**November 2004**
hep-th/0411163

# On the Coulomb Branch of a Marginal Deformation of $\mathcal{N}=4$ SUSY Yang-Mills

Nick Dorey

DAMTP, Centre for Mathematical Sciences
University of Cambridge, Wilberforce Road
Cambridge CB3 0WA, UK

and

Timothy J. Hollowood

Department of Physics, University of Wales Swansea
Singleton Park, Swansea, SA2 8PP, UK



**Abstract**

We determine the exact vacuum structure of a marginal deformation of $\mathcal{N}=4$ SUSY Yang-Mills with gauge group $U(N)$. The Coulomb branch of the theory consists of several sub-branches which are governed by complex curves of the form $\Sigma_{n_1}\cup\Sigma_{n_2}\cup\Sigma_{n_3}$ of genus $N=n_1+n_2+n_3$. Each sub-branch intersects with a family of Higgs and Confining branches permuted by $SL(2,\mathbf{Z})$ transformations. We determine the curve by solving a related matrix model in the planar limit according to the prescription of Dijkgraaf and Vafa, and also by explicit instanton calculations using a form of localization on the instanton moduli space. We find that $\Sigma_n$ coincides with the spectral curve of the $n$-body Ruijsenaars-Schneider system. Our results imply that the theory on each sub-branch is holomorphically equivalent to certain five-dimensional gauge theory with eight supercharges. This equivalence also implies the existence of novel confining branches in five dimensions.

# 1 Introduction

Four dimensional gauge theories with $\mathcal{N}=1$ supersymmetry are of both theoretical and phenomenological interest. The vacuum structure of these theories is governed by holomorphic observables which can often be computed exactly. Unlike other four-dimensional theories, one can determine the spectrum of massless particles and the breaking pattern of global and local symmetries even at strong coupling. In this paper we will determine the exact vacuum structure of an $\mathcal{N}=1$ model which arises as a marginal deformation of $\mathcal{N}=4$ SUSY Yang-Mills theory with gauge group $U(N)$ (we also discuss the $SU(N)$ theory).

In terms of $\mathcal{N}=1$ multiplets, the theory we will consider contains a single vector multiplet $V$ and three chiral multiplets $\Phi_i$ $i=1,2,3$ in the adjoint representation of the gauge group. The classical superpotential is,

$$\mathcal{W} = i\kappa \mathrm{Tr}_N \left[ e^{i\beta/2}\Phi_1\Phi_2\Phi_3 - e^{-i\beta/2}\Phi_1\Phi_3\Phi_2 \right] \ . \tag{1.1}$$

The $\mathcal{N}=4$ theory is recovered by setting $\beta=0$ and $\kappa=1$. In the following we will refer to the theory with $\beta \neq 0$ as the $\beta$-deformed theory. This theory is very special as the corresponding deformation of the $\mathcal{N}=4$ theory is exactly marginal and gives rise to a two-parameter family of $\mathcal{N}=1$ superconformal field theories [1].[1]

In the absence of scalar vacuum expectation values (VEVs), the $\beta$-deformed theory is exactly conformally invariant. However the theory also has branches of vacua in which conformal invariance is spontaneously broken. For generic values of the deformation parameter, the theory has Coulomb branches where the gauge group is broken down to its Cartan subalgebra. In addition, for certain special values, we find Higgs branches where the gauge group undergoes further breaking. In [3], it was argued that an exact S-duality, inherited from that of the undeformed $\mathcal{N}=4$ theory, implies the existence of dual branches where magnetic monopoles condense and external electric charges are confined. This phenomenon is familiar in the context of softly-broken $\mathcal{N}=2$ SUSY [5] and other $\mathcal{N}=1$ theories. However, there are some new features in this case. In particular, confinement occurs together with spontaneously broken conformal invariance, implying the existence of a massless composite dilaton. As discussed in [3, 4], the large-$N$ physics in these vacua exhibits further novel features like the appearance of additional dimensions, a relation to Little String Theory and a scaling limit where the worldsheet theory of the dual string is solvable. The new branches occur only occur at special values of the parameters where the theory is strongly-coupled and they are invisible in a classical analysis of the theory. Part of the motivation for the present investigation is to demonstrate the existence of these branches in a more direct way, without assuming S-duality.

[1]For early references on on the UV finiteness of this and other four-dimensional models see [2].

Initially, we will focus on the Coulomb branches mentioned above where the low-energy theory has gauge group $U(1)^N$. As in many other examples [5,6], the gauge couplings of this abelian effective theory are determined by the period matrix of a Riemann surface, or curve. However, the novel feature here is that the Coulomb branch has the form of the symmetric product $\mathrm{Sym}_N(\mathbf{C} \oplus \mathbf{C} \oplus \mathbf{C})$. This means that it actually consists of multiple intersecting sub-branches which classically can be described by VEVs of the scalar components of the 3 chiral fields:

$$\Phi_i = \mathrm{Diag}\big[x_1^{(i)}, \ldots, x_N^{(i)}\big] \ , \tag{1.2}$$

where for each $a$ only one of $x_a^{(i)}$ can be non-zero. Each sub-branch $\mathcal{C}_{n_1,n_2,n_3}$, with $\sum_i n_i = N$, is specified by the number $n_i$ of $x_a^{(i)}$ that are non-zero. One of our main results is that the curve on the sub-branch $\mathcal{C}_{n_1,n_2,n_3}$ actually has disjoint components consisting of three curves of genus $n_i$: $\Sigma_{n_1,n_2,n_3} = \Sigma_{n_1} \cup \Sigma_{n_2} \cup \Sigma_{n_3}$. Furthermore we identify each component $\Sigma_n$ as the spectral curve of the $n$-body Ruijsenaars-Schneider (RS) integrable system [26–28].

Following earlier work on massive versions of this model [7–9], we determine the curve (for the simplest case of the branch $\mathcal{C}_{N,0,0}$) using the correspondence between four dimensional $\mathcal{N} = 1$ theories and zero dimensional matrix models discovered by Dijkgraaf and Vafa [10]. We then confirm this result (and extend it to the general case) by calculating the instanton contributions to the couplings of the low-energy effective theory directly. The technical details of this latter calculation—albeit in a rather condensed form—are collected in Appendix B.

Our result for the curve $\Sigma_{n_1,n_2,n_3}$ allows us to identify the singular points where the Coulomb branch intersects with other Higgs and confining branches. The singularities occur at points on the Coulomb branch where states carrying electric and/or magnetic charges become massless. Although these states are not BPS in an $\mathcal{N} = 1$ theory, the monodromies of the period matrix allow us to identify the quantum numbers of the light states near each singular point. In this way we can explicitly check the existence of the required set of massless magnetic monopoles at the root of the confining branches discussed in [3,4]. Our results also provide further confirmation of the expected $SL(2,\mathbf{Z})$ duality inherited from the undeformed $\mathcal{N} = 4$ theory. In particular, this duality has a natural action on the moduli space of the curve $\Sigma_{n_1,n_2,n_3}$ which permutes the roots of Higgs and confining branches.

Remarkably, the spectral curve of the $N$-body RS model also governs the Coulomb branch of another $U(N)$ supersymmetric gauge theory [17]. This theory lives on the five dimensional spacetime $\mathbf{R}^{3,1} \times \mathbf{S}^1$. Its Lagrangian includes a bare mass term and has eight supercharges. Clearly this model is quite different from the $\beta$-deformed theory, which is four-dimensional, scale invariant and has only four supercharges. Despite these differences, the instanton calculations described in Appendix B prove rather directly that the Coulomb branches of the two theories are related by a simple holomorphic change of variables. We will also see that

the two theories have Higgs branches of the same dimensions for appropriate values of the parameters and they both have $SL(2,\mathbf{Z})$ dualities which act in the same way. Our conclusion is that there is a holomorphic equivalence between the two models. In other words, the two theories are equivalent at the level of $\mathcal{N}=1$ F-terms and differ only by $\mathcal{N}=1$ D-terms. This equivalence has some interesting consequences for the five-dimensional theory, which are discussed in the final Section of the paper.

The paper is organised as follows. In Section 2, we discuss the classical vacuum structure of the $\beta$-deformed theory. In Section 3, we investigate the form of perturbative and instanton corrections to the classical theory. In Section 4, we apply the method of Dijkgraaf and Vafa to obtain the conditions which define the curve $\Sigma_N$ on the branch $\mathcal{C}_{N,0,0}$. In Section 5, we present a new string theory derivation of the complex curve governing the quantum Coulomb branch of the five dimensional theory described above. In particular the string theory construction leads to the same defining conditions as those constraining $\Sigma_N$. We then show how the spectral curve of the RS system produces the unique solution of these conditions. In Section 6, we find explicit formulae for the low-energy abelian gauge couplings in the case of gauge group $U(2)$ and exhibit the roots of the Higgs and confining branches. Finally, in Section 7 we discuss some new features of the five-dimensional theory which can be inferred from its holomorphic equivalence to the $\beta$-deformed theory. Some calculational details are relegated to three Appendices. In particular, details of the instanton calculations described in the text are given in Appendix B. After this work was completed, we recieved the paper [18], which has some overlap with the results presented here.

# 2 Classical vacuum structure

In this Section we will discuss the classical vacuum structure of the $\beta$-deformed theory with gauge group $U(N)$. The F- and D-flatness conditions are,

$$[\Phi_1,\Phi_2]_\beta=[\Phi_2,\Phi_3]_\beta=[\Phi_3,\Phi_1]_\beta=0 \tag{2.1}$$

with

$$[\Phi_i,\Phi_j]_\beta=e^{i\beta/2}\Phi_i\Phi_j-e^{-i\beta/2}\Phi_j,\Phi_i \tag{2.2}$$

and

$$\sum_{i=1}^{3}[\Phi_i,\Phi_i^\dagger]=0\ , \tag{2.3}$$

respectively. For the $\mathcal{N}=4$ case, $\beta=0$, the deformed commutators appearing in the F-term constraint revert to ordinary ones. In this case the vacuum equations are solved by

diagonalizing each of the three complex scalars,

$$\langle \Phi_i \rangle = \mathrm{Diag}\left[x_1^{(i)}, x_2^{(i)}, \ldots, x_N^{(i)}\right] \tag{2.4}$$

The $3N^2$ complex eigenvalues $x_a^{(i)}$, for $i = 1,2,3$ and $a = 1,2,\ldots,N$, are unconstrained. After taking into account the Weyl group which permutes the eigenvalues, we recover the familiar Coulomb branch of the $\mathcal{N}=4$ theory. On this branch the $U(N)$ gauge symmetry is spontaneously broken down to its Cartan subalgebra $U(1)^N$ and the vacuum manifold is the symmetric product $\mathrm{Sym}_N\, \mathbf{C}^3$.

Introducing a generic, non-zero value of $\beta$ changes things considerably. The F-flatness conditions are no longer solved by arbitrary diagonal matrices (2.4). For each value of the Cartan index $a \in \{1,2,\ldots,N\}$, at most one of the three eigenvalues, $x_a^{(1)}$, $x_a^{(2)}$ and $x_a^{(3)}$, for each $a$, can be non-zero. To this end, we define the three subgroups

$$\mathfrak{I}_i = \left\{ a \;\middle|\; a \in \{1,\ldots,N\}\ ,\ x_a^{(i)} \neq 0 \right\}\ , \quad i = 1,2,3\ ; \tag{2.5}$$

hence, $\mathfrak{I}_1 \cup \mathfrak{I}_2 \cup \mathfrak{I}_3 = \{1,\ldots,N\}$ and $\mathfrak{I}_i \cap \mathfrak{I}_j = \emptyset$ for $i \neq j$. We also find it convenient to define for each $a = 1,\ldots,N$

$$x_a = x_a^{(i)} \text{ when } a \in \mathfrak{I}_i\ . \tag{2.6}$$

In the simplest case of gauge group $U(1)$, the Coulomb branch of the $\mathcal{N}=4$ theory, which has three complex dimensions, is partially lifted leaving three complex lines which intersect at the origin. For $G = U(N)$ with $N > 1$, the Coulomb branch is formed by taking an $N$-fold symmetric product in the usual way to give $\mathrm{Sym}_N(\mathbf{C}\oplus\mathbf{C}\oplus\mathbf{C})$. So in the $\beta$-deformed theory there are multiple Coulomb branches rather than a single branch as in an $\mathcal{N}=2$ theory. Up to gauge invariance, the inequivalent branches are labelled as $\mathcal{C}_{n_1,n_2,n_3}$, where $n_i = \dim \mathfrak{I}_i$. Often we will pay particular attention to the three branches $\mathcal{C}_1 \equiv \mathcal{C}_{N,0,0}$, etc., on which only one of the one of the three complex adjoint scalars is non-zero.

The Lagrangian corresponding to the superpotential (1.1) has a $U(1)^3 \simeq U(1)_R^{(1)} \times U(1)_R^{(2)} \times U(1)_R^{(3)}$ R-symmetry, where each complex scalar field $\Phi_i$ is charged under $U(1)_R^{(i)}$ and neutral under the other factors. At a generic point on the Coulomb branch $\mathcal{C}_{n_1,n_2,n_3}$ the complete R-symmetry group is is spontaneously broken and the gauge group is spontaneously broken to $U(1)^N$. The branch can also be parametrized in terms of the gauge invariant moduli,

$$u_n^{(i)} = \frac{1}{N}\langle \mathrm{Tr}\, \phi_i^n \rangle = \frac{1}{n_i}\sum_{a\in\mathfrak{I}_i} x_a^n\ . \tag{2.7}$$

for $n = 1,2,\ldots,n_i$.

As the unbroken gauge group is $U(1)^N$, the massless fields at a generic point on the branch includes $N$ massless photons corresponding to the diagonal elements $A_{an} \equiv (A_n)_{aa}$

of the gauge field and their gluino superpartners $\lambda_{a\alpha} \equiv (\lambda_\alpha)_{aa}$ which make up $N$ abelian vector multiplets of $\mathcal{N}=1$ SUSY. We denote the corresponding field-strength superfields $W_{a\alpha}$, $a=1,2,\ldots,N$. There are also $N$ massless chiral multiplets, neutral under $U(1)^N$, corresponding to fluctuations of the eigenvalues $x_a$. These fields are associated to the subset of the diagonal elements of $\Phi_i=(\phi_i,\psi_{i\alpha})$; namely, $\phi_a=(\phi_i)_{aa}$ and $\psi_{a\alpha}=(\psi_{i\alpha})_{aa}$ for $a\in\mathfrak{I}_i$.

In addition to the massless states described above the full classical spectrum includes states which have masses due to the Higgs mechanism. Such states arise from the non-diagonal elements of the $U(N)$ vector multiplet $V$ and also from the non-diagonal plus some of the diagonal elements of the three chiral multiplets $\Phi_i$, $i=1,2,3$. First of all, for each pair $a\in\mathfrak{I}_i$ and $b\in\mathfrak{I}_j$, $i\neq j$, there is a vector multiplet and three chiral multiplets of mass squared $|x_a|^2+|x_b|^2$. For $a\neq b$ but when $a,b\in\mathfrak{I}_i$ there is a vector multiplet and three chiral multiplets of masses $M^v_{ab}=|(Z_v)_{ab}|$ and $M^j_{ab}=|(Z_j)_{ab}|$, respectively, where

$$
\begin{aligned}
(Z_v)_{ab} &= (Z_i)_{ab} = x_a - x_b \ , \\
(Z_{i+1})_{ab} &= e^{i\beta/2}x_a - e^{-i\beta/2}x_b \ , \\
(Z_{i-1})_{ab} &= e^{-i\beta/2}x_a - e^{i\beta/2}x_b \ .
\end{aligned}
\tag{2.8}
$$

(The labels $i$, etc., are to be understood as defined modulo 3.) Finally of each $a\in\mathfrak{I}_i$ there are there are two additional massive chiral multiplets of mass $M^{i\pm1}_{aa}=2|x_a|\sin(\beta/2)$ coming from the diagonal elements of $\Phi_{i\pm1}$.

For generic values of the deformation parameter the massless multiplets identified above correspond to the vanishing diagonal elements $(Z_v)_{aa}$ and $(Z_i)_{aa}$, $a\in\mathfrak{I}_i$. However, for special values of the eigenvalues $x_a$, additional massless states appear. In some cases, these extra massless states indicate the freedom to move off along new branches. For example, for $a\in\mathfrak{I}_i$ there are always additional massless chiral multiplets $(\Phi_j)_{aa}$, $j\neq i$ on submanifolds where $x_a=0$. These are points where the three different Coulomb branches $\mathfrak{C}_{p_1+1,p_2,p_3}$, $\mathfrak{C}_{p_1,p_2+1,p_3}$ and $\mathfrak{C}_{p_1,p_2,p_3+1}$ intersect. In these cases the new massless states are uncharged. As usual there are also subspaces on the classical Coulomb branch where a non-abelian subgroup of the gauge group is restored. For example, an $SU(2)$ subgroup is restored when $x_a=x_b$, for $a,b\in\mathfrak{I}_i$. The resulting non-abelian low-energy theory is typically asymptotically free and runs to strong coupling in the IR invalidating a classical analysis.

Finally, there are also points where we find additional charged massless states without the restoration of non-abelian gauge symmetry. In particular, this occurs on the submanifolds where $e^{i\beta/2}x_a=e^{-i\beta/2}x_b$, for $a,b\in\mathfrak{I}_i$. In these cases the resulting low-energy theory is typically IR free and we will below that these singular submanifolds persist in the quantum theory. Further for certain non-generic rational values of $\beta$, the number of massless states is large enough to result in a new branch on which these degrees of freedom condense further breaking the gauge group.

**Example: the $U(2)$ theory**

For simplicity, let us consider the $U(2)$ theory and its Coulomb branch $\mathcal{C}_{2,0,0}$ parameterized by eigenvalues $x_1$ and $x_2$. New massless states appear on the one-dimensional submanifolds defined by $x_1 = \exp(\pm i\beta)x_2$. The gauge invariant version of this condition (for $\beta \neq \pi/2$) is,

$$u_1^2 = \frac{\cos^2(\beta/2)}{\cos(\beta)}\, u_2 \ . \tag{2.9}$$

For either root, we find light fields charged under $U(1)_1 \times U(1)_2$. For example, near the submanifold corresponding to $x_1 = \exp(-i\beta)x_2$, we have two light chiral superfields $Q = (\Phi_2)_{12}$ and $\tilde{Q} = (\Phi_3)_{21}$ with charges $(+1,-1)$ and $(-1,+1)$. This matter content is equivalent to a single $\mathcal{N} = 2$ hypermultiplet and we will sometimes use this language in the following (although the theory only has $\mathcal{N} = 1$ SUSY). Apart from the gauge couplings, the effective theory of the light fields has a superpotential,

$$W_{\text{eff}} = \left(e^{-i\beta/2}x_1 - e^{i\beta/2}x_2\right) Q\tilde{Q} \tag{2.10}$$

a similar effective theory arises near the other root, corresponding to the submanifold $x_1 = \exp(-i\beta)x_2$.

An important special case is when $\beta = \pi$ where the two roots of (2.9) coincide at $u_1 = 0$ or $x_1 + x_2 = 0$. Near this point in the moduli space we now find two light hypermultiplets. Equivalently we find two light chiral superfields, $Q_1 = (\Phi_2)_{12}$ and $Q_2 = (\Phi_3)_{12}$, with charges $(+1,-1)$, under $U(1) \times U(1)$ and two more, denoted $\tilde{Q}_1 = (\Phi_2)_{21}$ and $\tilde{Q}_2 = (\Phi_3)_{21}$ with charges $(-1,+1)$. In addition to the gauge couplings, the low energy effective theory near this point has superpotential

$$W = (x_1 + x_2)\left(Q_1\tilde{Q}_1 + Q_2\tilde{Q}_2\right) \tag{2.11}$$

A key feature of this effective theory is the existence of a new Higgs branch on which the massless charged states condense. The new branch appears for $x_1 + x_2 = 0$ and allows non-zero values for the charged fields subject to the F- and D-term conditions,

$$\begin{aligned} Q_1\tilde{Q}_1 + Q_2\tilde{Q}_2 &= 0 \\ |Q_1|^2 - |\tilde{Q}_1|^2 + |Q_2|^2 - |\tilde{Q}_2|^2 &= 0 \ . \end{aligned} \tag{2.12}$$

By fixing the $U(1) \times U(1)$ gauge symmetry and using the above relations, we may eliminate $\tilde{Q}_1$ and $\tilde{Q}_2$, leaving a three complex dimensional branch of solutions parametrized by $Q_1$, $Q_2$ and $x_1 + x_2$. When either $Q_1$ or $Q_2$ is non-zero the $U(1) \times U(1)$ gauge symmetry of the Coulomb branch is broken down to the diagonal $U(1)$.

It is not hard to find the corresponding branch in the full $\beta$-deformed theory. When $\beta = \pi$, the deformed commutator $[\Phi_i, \Phi_j]_\beta$ appearing in (1.1) becomes the anti-commutator $\{\Phi_i, \Phi_j\}$ and we can solve the F- and D-flatness conditions by setting,

$$\langle \Phi_1 \rangle = \alpha_1 \tau_3 \qquad \langle \Phi_2 \rangle = \alpha_2 \tau_1 \qquad \langle \Phi_3 \rangle = \alpha_3 \tau_2 \tag{2.13}$$

wherend $\tau_i$ are the Pauli matrices and $\alpha_1$, $\alpha_2$ and $\alpha_3$ are arbitrary complex numbers. When two or more of the $\alpha_i$ are non-zero the $U(2)$ gauge group is broken to its central $U(1)$. This Higgs branch intersects the Coulomb branch $\mathcal{C}_{(2,0,0)}$ at $u_1 = 0$ when $\alpha_2 = \alpha_3 = 0$. The massless modes on the Higgs branch include three scalars corresponding to fluctuations of the moduli $\alpha_i$, $i = 1, 2, 3$ and the photon of the central $U(1)$. Each of these bosonic fields is paired with a massless Weyl fermion by the unbroken $\mathcal{N} = 1$ supersymmetry. These fields are free at low energies and the effective action is precisely that of an $\mathcal{N} = 4$ supersymmetric gauge theory with gauge group $U(1)$. The complexified gauge coupling of the low-energy theory is related to that of the original theory as $\tilde{\tau} = 2\tau$.

We will now give a brief (and incomplete) discussion of the Higgs branches which appear in the $U(N)$ theory for arbitrary $N$. As in the $N = 2$ case, the theory with $\beta = 2\pi/N$ has a Higgs branch where $U(N)$ is broken broken to its central $U(1)$. The root occurs at a point on the Coulomb branch $\mathcal{C}_1$ where the eigenvalues of $\Phi_1$ take the values $x_a = \alpha_1 \exp(2\pi i a/N)$ for $a = 1, 2, \ldots, N$. As above $\alpha_1$ is an arbitrary non-zero complex number. At this point we find $N$ massless chiral superfields $Q_a$, for $a = 1, 2, \ldots, N$ which carry charges,

$$\begin{array}{c} (+1, -1, 0, \ldots, 0) \\ (0, +1, -1, \ldots, 0) \\ \cdots\cdots\cdots\cdots \\ (0, 0 \ldots, +1, -1) \\ (-1, 0, \ldots, 0, +1) \end{array} \tag{2.14}$$

under the unbroken $U(1)^N$ gauge symmetry, as well as charge-conjugate degrees of freedom contained in chiral superfields $\tilde{Q}_a$ with the opposite charges. The effective theory has also has a superpotential which is trilinear in $x_a$, $Q_a$ and $\tilde{Q}_a$. The matter content and interactions are essentially those of an $\mathcal{N} = 2$ quiver theory with gauge group $U(1)^N$ corresponding to the Dynkin diagram of the $A_{N-1}$ Lie algebra. The latter theory is known to have a Higgs branch where $U(1)^N$ is broken to its diagonal $U(1)$ subgroup.

The corresponding Higgs branch of the full $\beta$-deformed theory has scalar expectation values,

$$\langle \Phi_1 \rangle = \alpha_1 U_{(N)} \qquad \langle \Phi_2 \rangle = \alpha_2 V_{(N)} \qquad \langle \Phi_3 \rangle = \alpha_3 W_{(N)} \tag{2.15}$$

where $U_{(N)}$ and $V_{(N)}$ are the $N \times N$ "clock" and "shift" matrices, $(U_{(N)})_{ab} = \delta_{ab} \exp(2\pi i a/N)$ and

$$\left(V_{(N)}\right)_{ab} = \begin{cases} 1 & \text{if } b = a + 1, \mod N \\ 0 & \text{otherwise} \end{cases} \tag{2.16}$$

and $W_N = V^\dagger_{(N)} U^\dagger_{(N)}$. Here $\alpha_1$, $\alpha_2$ and $\alpha_3$ are complex numbers.

The $U(N)$ theory can also have more general Higgs branches with a larger unbroken gauge group. These occur when the rank $N$ has a non-trivial divisor. Thus we have $N = mn$ for some integers $m$ and $n$. If the deformation parameter takes the value $\beta = 2\pi/n$ we find a $3m$ complex parameter branch

$$\langle \Phi_1 \rangle = \Lambda^{(1)} \otimes U_{(n)} \ , \qquad \langle \Phi_2 \rangle = \Lambda^{(2)} \otimes V_{(n)} \ , \qquad \langle \Phi_3 \rangle = \Lambda^{(3)} \otimes W_{(n)} \ , \tag{2.17}$$

where $\Lambda^{(i)}$ $i = 1, 2, 3$ are three arbitrary diagonal $m \times m$ matrices. At a generic point on this branch the unbroken gauge symmetry is $U(1)^m$. A special case occurs when each $\Lambda^{(i)}$ is proportional to the $m \times m$ unit matrix. In this three complex parameter subspace the unbroken gauge group is enhanced to $U(m)$.

In addition to the Coulomb and Higgs branches described above, there are also mixed branches which can be constructed in the obvious way when $e^{ip\beta} = 1$ for $p < N$.

To close this Section we will discuss the case of gauge group $SU(N)$. At the classical level, the relation between the $U(N)$ and $SU(N)$ theories defined by the superpotential (1.1) is non-trivial. Apart from the central photon and its $\mathcal{N} = 1$ superpartner, the $U(N)$ theory also contains three chiral superfields $a_i = \mathrm{Tr}_N \Phi_i$ for $i = 1, 2, 3$, which are not present in the $SU(N)$ theory. While the central $U(1)$ vector multiplet is completely decoupled from the $SU(N)$ degrees of freedom, the three chiral multiplets do not (for $\beta \neq 0$). All the branches of the $U(N)$ theory discussed above are also present in the $SU(N)$ theory, although, in some cases, their complex dimension is reduced by the traceless condition $a_i = 0$.

# 3 The Quantum Theory

The superpotential (1.1) corresponds to an exactly marginal deformation of $\mathcal{N} = 4$ SUSY Yang-Mills with gauge group $SU(N)$ [1]. The coupled $\beta$-functions for the couplings $\tau$, $\beta$ and $\kappa$ vanish on a two (complex) dimensional surface in the parameter space. This surface includes the $\mathcal{N} = 4$ line, $\beta = 0$, $\kappa = 1$ with $\tau$ arbitrary. Away from this line the critical surface is specified as $\kappa = \kappa_{cr}[\tau, \beta]$ however the explicit form of $\kappa_{cr}$ is unknown beyond one-loop. As mentioned above, the $U(N)$ theory is classically equivalent to the $SU(N)$ theory with additional couplings to the trace chiral multiplets $a_i$, for $i = 1, 2, 3$. In the quantum theory, these additional couplings are actually IR free and thus the trace fields decouple from the $SU(N)$ degrees of freedom in the IR [2]. At the quantum level, therefore, the $U(N)$ theory at the origin contains the $SU(N)$ conformal theory plus some additional free fields.

[2]The authors acknowledge a useful discussion with Ofer Aharony on this point.

The vanishing $\beta$-functions imply that the deformed theory has exact $\mathcal{N}=1$ superconformal invariance and that there are no chiral anomalies. Thus the $U(1)^3$ R-symmetry of the classical theory persists in the full quantum theory although it is broken spontaneously on the Coulomb branches. These symmetries prevent the Coulomb branch from being lifted by quantum effects.[3] Thus on a given Coulomb branch, we find a moduli space of vacua with $N$ massless $U(1)$ vector multiplets $W_{a\alpha}=(\lambda_{a\alpha},A_{an})$, which we can think of the the diagonal components of the the vector multiplet of the microscopic theory, and $N$ neutral chiral multiplets $\Phi_a=(\phi_a,\psi_{a\alpha})$ which we can think of as the diagonal elements $(\Phi_i)_{aa}$ for $a\in\mathfrak{I}_i$. Although no superpotential can be generated, we expect the kinetic terms of the massless fields to recieve quantum corrections. As we only have $\mathcal{N}=1$ SUSY the kinetic terms for the scalars correspond to D-terms which are relatively unconstrained. In contrast the exact effective action for the massless gauge fields is an F-term of the form

$$\mathcal{L}_{\rm eff}=\frac{1}{8\pi}\,{\rm Im}\left[\int d^2\theta\sum_{ab=1}^{N}\tau_{ab}(\Phi_c)\,W_\alpha^{\ a}W^{\alpha b}\right]\ . \tag{3.1}$$

The effective gauge couplings and vacuum angles are encoded in the complex $N\times N$ matrix $\tau_{ab}$, which depends holomorphically on the $N$ effective chiral superfields $\Phi_a$, $a=1,2,\ldots N$. In particular, the coupling constants of the abelian gauge fields at a point on the Coulomb branch are $\tau_{ab}(x_c)$.

At the classical level the matrix of effective couplings is simply $\tau^{\rm cl}_{ab}=\delta_{ab}\tau$. Holomorphy constrains the possible quantum corrections precisely as in an $\mathcal{N}=2$ theory. At the perturbative level, only one-loop corrections are allowed while beyond perturbation theory instanton contributions are allowed of arbitrary charge. These latter contributions are at leading order in the semi-classical approximation which is valid for large VEVs. Schematically,

$$\tau_{ab}(x_c)=\tau\delta_{ab}+\tau_{ab}^{\text{1-loop}}(x_c)+\sum_{k=1}^{\infty}\tau_{ab}^{k-inst.}(x_c)e^{2\pi ik\tau}\ . \tag{3.2}$$

In the next subsections, we will examine the pertubative and non-perturbative corrections in turn.

## 3.1 Perturbation theory

It is straightforward to calculate the effective couplings at one-loop in perturbation theory. The result only depends on the masses, spins and abelian charges of the states that can

[3] Essentially, a term with the correct R-charge and dimension in the effective superpotential would have to be proportional to $\Phi_1\Phi_2\Phi_3$ with some contraction of the group indices. No such term would lift the Coulomb branch.

propagate in the loop. It is helpful to organize the massive states into multiplets of $\mathcal{N}=1$ supersymmetry.

The result is that $\tau_{ab}^{\text{1-loop}}=0$ for $a=b$ and also for $a\in\mathfrak{I}_i$ and $b\in\mathfrak{I}_j$, when $i\neq j$. For $a,b\in\mathfrak{I}_i$ (but $a\neq b$) we have the non-vanishing contribution

$$\tau_{ab}^{1-\text{loop}}=f_{ab}-\delta_{ab}\Big(\sum_{c\neq a}f_{ac}\Big)\ , \tag{3.3}$$

where

$$f_{ab}=\frac{i}{2\pi}\log\left[\frac{(Z_v)^3_{ab}}{(Z_1)_{ab}\,(Z_2)_{ab}\,(Z_3)_{ab}}\right]\ . \tag{3.4}$$

This result reflects the contributions of each of the massive states identified in (2.8) as virtual particles running around the loop. This formula also exhibits the conformal R-symmetry properties of the theory which imply that $\tau_{ab}$ is invariant under the transformation $x_a\to\lambda x_a$ for $a=1,2,\ldots,N$ and $\lambda$ is any complex number.

Apart from holomorphy, symmetries and the perturbative limit, there are other constraints on the exact form of the effective gauge couplings. For example, the unitarity of the low energy theory requires $\mathrm{Im}[\tau_{ab}]\geq 0$. The low-energy theory on the Coulomb branch $\mathcal{C}_1$ is also invariant under $Sp(2N,\mathbf{Z})$ electric-magnetic duality transformations acting on the low-energy couplings. This means that, in general, $\tau_{ab}$ will not be a single-valued function on the moduli space, but can exhibit non-trivial $Sp(2N,\mathbf{Z})$ monodromies around singular points (or, more generally, singular submanifolds). Such singular points occur where charged degrees of freedom become massless and the monodromies reflect the one-loop $\beta$-function of the effective theory of the light degrees of freedom near the singular point. An example of this behaviour is already evident in the perturbative result (3.3). For simplicity we focus on the case $N=2$.

On the Coulomb branch $\mathcal{C}_1$ of the $U(2)$ theory the gauge symmetry is broken down to $U(1)_1\times U(1)_2$, where the two $U(1)$ factors, generated by $Q_1$ and $Q_2$ respectively, correspond to the two diagonal elements of the $U(2)$ gauge field. It is convenient to change basis to $U(1)_{\text{even}}\times U(1)_{\text{odd}}$ generated by $(Q_1\pm Q_2)/2$ respectively. In terms of the decomposition,

$$U(2)\simeq\frac{U(1)\times SU(2)}{\mathbf{Z}_2} \tag{3.5}$$

$U(1)_{\text{even}}$ corresponds to the center of $U(2)$ and $U(1)_{\text{odd}}$ corresponds to the Cartan subalgebra of $SU(2)$. They are even and odd respectively under the Weyl group of $SU(2)$ which permutes $U(1)_1$ and $U(1)_2$.

As the gauge boson of $U(1)_{\text{even}}$ is decoupled the matrix of abelian couplings is diagonal in this basis: $\tau_{ab}=\mathrm{diag}(\tau_{\text{even}},\tau_{\text{odd}})$. Including classical and one-loop effects we find, $\tau_{\text{even}}=\tau$

and

$$\tau_{\mathrm{odd}} = \tau + \frac{1}{\pi i}\log\left[\frac{(x_1-x_2)^2}{(e^{i\beta/2}x_1 - e^{-i\beta/2}x_2)(e^{-i\beta/2}x_1 - e^{i\beta/2}x_2)}\right] \ . \tag{3.6}$$

Clearly this expression exhibits a logarithmic singularity on the submanifolds $x_1 = \exp(\pm i\beta)x_2$ where new massless states appear in the classical theory. It is convenient to introduce the gauge-invariant modulus,

$$\varphi = \frac{u_1}{\sqrt{2u_1^2 - u_2}} = \frac{x_1 + x_2}{2\sqrt{x_1 x_2}} \ . \tag{3.7}$$

The singular submanifolds lie at $\varphi = \pm\cos(\beta/2)$. the leading behaviour of $\tau_{\mathrm{odd}}$ near the point $\varphi = \cos(\beta/2)$ is

$$\tau_{\mathrm{odd}} \sim -\frac{1}{\pi i}\log\big(\varphi - \cos(\beta/2)\big) \ . \tag{3.8}$$

Thus we see that the effective coupling undergoes a monodromy

$$\mathcal{M}_1 : \qquad \tau_{\mathrm{odd}} \to \tau_{\mathrm{odd}} - 2 \ , \tag{3.9}$$

as we traverse a small circle in the complex $\rho$ plane enclosing the point $\varphi = \cos(\beta/2)$ in an anti-clockwise direction.

In order to have a globally consistent description of the theory with $\mathrm{Im}[\tau_{ab}] > 0$ everywhere we must find additional singular submanifolds with associated $Sp(2N, \mathbf{Z})$ monodromies which do not commute with $\mathcal{M}_1$. These conditions can be satisfied by identifying $\tau_{ab}$ with the period matrix of an appropriate family of complex curves of genus $N$, just as in an $\mathcal{N} = 2$ theory. In the next section, we will determine this curve explicitly.

## 3.2 Instanton effects

The instanton contributions to the couplings $\tau_{ab}$ in the low-energy effective action (3.2) can be calculated in much the same way as in an $\mathcal{N} = 2$ theory. Recall that the massless fields correspond to $N$ abelian vector multiplets of $\mathcal{N} = 1$ supersymmetry and $N$ neutral chiral multiplets. The fermionic components of these multiplets are $\lambda_{a\alpha}$, the gluinos, and $\psi_{a\alpha}$. (Recall that $\psi_{a\alpha}$ and its bosonic partner $\psi_a$, for $a \in \mathfrak{I}_i$, come from the corresponding diagonal component of the chiral multiplet $\Phi_i$.) As in an $\mathcal{N} = 2$ theory, the instanton contribtion to $\tau_{ab}$ can be extracted from the universal long distance behaviour of the *anti*-fermion correlator

$$\begin{aligned}&\langle \bar\lambda_a^{\dot\alpha}(x^{(1)})\bar\lambda_b^{\dot\beta}(x^{(2)})\bar\psi_c^{\dot\gamma}(x^{(3)})\bar\psi_d^{\dot\delta}(x^{(4)})\rangle \\ &= \int d^4X \bar S^{\dot\alpha\alpha}(x^{(1)} - X)\bar S^{\dot\beta}_{\alpha}(x^{(2)} - X)\bar S^{\dot\gamma\beta}(x^{(3)} - X)\bar S^{\dot\delta}_{\beta}(x^{(4)} - X)\frac{\tau_{ab}}{\partial\phi_c\partial\phi_d}\Big|_{\phi_a = x_a} + \cdots \ ,\end{aligned} \tag{3.10}$$

where

$$\bar{S}^{\dot{\alpha}\alpha}(x) = \frac{1}{4\pi^2}\,\bar{\partial\!\!\!/}^{\dot{\alpha}\alpha}\Big(\frac{1}{x^2}\Big) \tag{3.11}$$

is the free anti-Weyl spinor propagator.

In order to calculate the instanton contribution to this correlator one has to insert the leading-order semi-classical expressions for the anti-fermions in the instanton background into the measure for integrating over the supersymmetric multi-instanton moduli space $\mathfrak{M}_k$.[4] Note that in our theory the zero-mode structure is identical to the $\mathcal{N}=4$ theory and the measure is schematically of the form

$$\int_{\mathfrak{M}_k} \boldsymbol{\omega}^{(\mathcal{N}=4)} e^{-\tilde{S}_k^{(\beta)}} \tag{3.12}$$

where $\boldsymbol{\omega}^{(\mathcal{N}=4)}$ is the volume form for integrating over super moduli space of instantons in the $\mathcal{N}=4$ theory. The $\beta$-deformation appears explicitly in $\tilde{S}_k^{(\beta)}$, the instanton effective action which depends on the collective coordinates of the instanton. The construction of this action is outlined in Appendix B. In comparison with the $\mathcal{N}=4$ theory, an instanton configuration only has two supersymmetric zero-modes. This means with the $\beta$-deformation, the instanton effective action is only independent of two of the Grassmann collective coordinates $\xi_\alpha$ associated to the supersymmetric zero modes (coming from the gluino).

In the instanton backgound, the long distance behaviour of the anti-fermions has the universal form

$$\begin{aligned}\bar{\psi}_a^{\dot{\alpha}}(x) &= \bar{S}^{\dot{\alpha}\alpha}(x-X)\Theta_{\alpha a} + \cdots\ ,\\ \bar{\lambda}_a^{\dot{\alpha}}(x) &= \bar{S}^{\dot{\alpha}\alpha}(x-X)\Xi_{\alpha a} + \cdots\ ,\end{aligned} \tag{3.13}$$

where $\Theta_{\alpha a}$ and $\Xi_{\alpha a}$ are functions of the collective coordinates of the super-instanton that are linear in the Grassmann ones. In the above, $X$ is the centre of the instanton which is identified with the integration variable in (3.10). The two supersymmetric collective coordinates $\xi_\alpha$ only appear $\Theta_{\alpha a}$ and the integrals over these variables must therefore be saturated by the two $\bar{\psi}$ insertions. The precise form of the relevant terms is

$$\Theta_{a\alpha} = \xi_\alpha \frac{\partial \tilde{S}^{(\beta)}}{\partial x_a} + \cdots\ , \tag{3.14}$$

and therefore

$$\tau_{ab}^{\text{inst.}} = \sum_{k=1}^{\infty} e^{2\pi i k\tau} \int_{\widehat{\mathfrak{M}}_k} \boldsymbol{\omega}^{(\beta)} e^{-\tilde{S}^{(\beta)}}\, \Xi_a^\alpha\, \Xi_{b\alpha}\ , \tag{3.15}$$

[4] We use throughout the notation of [25].

where $\boldsymbol{\omega}^{(\beta)}$ is the integral over the $\mathcal{N}=1$ centred instanton moduli space $\widehat{\mathfrak{M}}_k$ which is, schematically, $\boldsymbol{\omega}^{(\mathcal{N}=4)}/d^4X\, d^2\xi$.

At the one instanton level we can prove that $\tau_{ab}\neq 0$ only if $a,b\in\mathfrak{I}_i$. the argument relies on the fact that $\Xi_{a\alpha}$ is proportional to one of the Grassmann collective coordinates associated to the zero-modes of the $\mathcal{N}=4$ theory that are lifted by the $\beta$-deformations. We denote these as $\xi^i_\alpha$, for $i\in\{1,2,3\}$. More specifically, for $a\in\mathfrak{I}_i$, the exact long distance behaviour is captured by the exact expression

$$\Xi_{a\alpha}=\xi^i_\alpha\frac{\partial\tilde{S}^{(\beta)}}{\partial x_a}\ . \tag{3.16}$$

It follows therefore that

$$\tau^{\text{1-inst.}}_{ab}=\frac{\partial^2 F}{\partial x_a\partial x_b}\ , \tag{3.17}$$

for some function $F$. Note that $F$ is like the prepotential of an $\mathcal{N}=2$ theory except that in the $\mathcal{N}=1$ context there is no need for (3.17) to be true for all instanton number. It is given by an integral over the moduli space quotiented by $\{X_n,\xi_\alpha,\xi^i_\alpha\}$. It immediately follows from the existence of $F$ that $\tau_{ab}=0$ for $a\in\mathfrak{I}_i$ and $b\in\mathfrak{I}_j$, when $i\neq j$. The reason is that $\tau_{ab}$ is uncharged under the three R-symmetries $U(1)^{(i)}_R$, $i=1,2,3$. Then, since $x_a$ has R-charge $+2$ under $U(1)^{(i)}_R$, for $a\in\mathfrak{I}_i$, and is uncharged under the remaining two R-symmetries, it must be that $F=F_1+F_2+F_3$, where $F_i$ depends only on $\{x_a,a\in\mathfrak{I}_i\}$, and $F_i$ must have the same R-charge as $x_a^2$, $a\in\mathfrak{I}_i$. Consequently, $\tau_{ab}$ decomposes into three separate blocks at the one-instanton level.

It turns out that this block structure generalizes to higher instanton numbers so that the matrix of couplings is block-diagonal to all orders in the instanton expansion. The proof of this relies on detailed localization arguments in the instanton calculus and we relegate them to Appendix B. So in an exact sense each Coulomb branch $\mathcal{C}_{n_1,n_2,n_3}$ decomposes into a direct sum of three blocks where the couplings within each block only depend on the VEVs associated to that block. In addition, when one of the VEVs, say $x_a$ goes to zero, the corresponding couplings $\tau_{ab}$ and $\tau_{ba}$, for $b\neq a$ go to zero. These subspaces describe the intersections of different Coulomb branches.

# 4 The Seiberg-Witten Curve from the Dijkgraaf-Vafa Matrix Integral

The Dijkgraaf-Vafa matrix model gives a way of computing an effective superpotential in $\mathcal{N}=1$ SYM in terms of the glueball superfields [10–12]. In our case, there is a moduli space

of vacua and no superpotential, however, as described originally in [13], the matrix model technique can be used to re-construct a Coulomb branch by choosing a suitable deformation which allows one to lift the degeneracy of the Coulomb branch in a way that allows one to probe an arbitrary point. This technique is only possible on the Coulomb branches $\mathcal{C}_i$ where all the VEV reside in one of the chiral fields. For definitness we choose $\mathcal{C}_1$.

In order to probe $\mathcal{C}_1$ we deform the theory by a superpotential for $\Phi_1$ of the form $\mathrm{Tr}\, V(\Phi_1)$ where

$$V'(x) = \mu \prod_{a=1}^{N} (x - \xi_a) \ . \tag{4.1}$$

Classically, the potential lifts $\mathcal{C}_1$ and leaves an isolated vacuum at $x_a = \xi_a$.[5] We now briefly describe how to apply the matrix model of Dijkgraaf-Vafa to find the Seiberg-Witten curve on $\mathcal{C}_1$.

The matrix model involves three matrices, one for each of the chiral superfields $\Phi_i$ (see [7, 14, 15] for a discussion of these kinds of matrix models). The matrix model partition function involves the superpotential of the parent field theory:

$$Z = \int \prod_{i=1}^{3} d\Phi_i \, \exp -g_s^{-1} \mathrm{Tr} \left( i\kappa \Phi_1 [\Phi_2, \Phi_3]_\beta + V(\Phi_1) \right) \ . \tag{4.2}$$

Note that we use the same notation for the matrices as for their associated chiral superfields. The first thing to do involves integrating out $\Phi_{2,3}$ by choosing a suitable contour on which the integrals are well-defined. Then one can go to eigenvalue basis for $\Phi_1$:

$$Z \sim \int \prod_{a=1}^{N} dx_a \, \frac{\prod_{a\neq b}(x_a - x_b)}{\prod_{ab}(e^{i\beta/2} x_a - e^{-i\beta/2} x_b)} \exp -g_s^{-1} \sum_a V(x_a) \ . \tag{4.3}$$

Note that the denominator in the above comes from integrating out $\Phi_2$ and $\Phi_3$ while the numerator is the famous Van der Monde determinant arising from going to the eigenvalue basis for $\Phi_1$.

Recall that the we want to study the vacuum of the parent theory where classically $x_a = \xi_a$, $a = 1, \ldots, N$. In the matrix model, one develops a saddle-point expand around this classical solution. This is achieved by replacing the size of the matrices by $\hat{N}$ and then by taking $\hat{N} \to \infty$ with $g_s \to 0$ keeping $S = g_s \hat{N}$ fixed. In more detail, we take the saddle-point around the classical critical point with $\hat{N}_a$ eigenvalues at $\xi_a$. Obviously $\sum_{a=1}^{N} \hat{N}_a = \hat{N}$ and

[5] There are of course additional vacua, but we focus on the one where one eigenvalue of $\Phi_1$ is associated to each of the $N$ minima of $V(x)$.

then we take each $\hat{N}_a \to \infty$ independently keeping the $N$ quantities

$$S_a = g_s \hat{N}_a \ , \qquad S = \sum_{a=1}^{N} S_a \ , \tag{4.4}$$

fixed. We emphasize that $\hat{N}_a$ are not the physical degeneracies which are equal to one.

The saddle-point equation which follows from (4.3) is

$$g_s \Big[ 2 \sum_{b(\neq a)} \frac{1}{x_a - x_b} - \sum_b \frac{1}{x_a - e^{i\beta} x_b} - \sum_b \frac{1}{x_a - e^{-i\beta} x_b} \Big] = V'(x_a) \ . \tag{4.5}$$

The terms on the left-hand side correspond to quantum effects which modify the classical saddle-point solution. In the large $\hat{N}$ limit, we can describe the eigenvalues with a density $\rho(x)$ in the complex eigenvalue $x$-plane. Experience suggests that $\rho(x)$ has support along $N$ open contours $\mathcal{C}_a$ in the neighbourhood of each $\xi_a$. We will normalize the density via

$$\sum_a \int_{\mathcal{C}_a} \rho(x)\, dx = 1 \ . \tag{4.6}$$

A key related quantity is the *resolvent* $\omega(x)$ which is an analytic function on the $x$-plane with $N$ branch cuts along each $\mathcal{C}_a$ defined in terms of the density via

$$\omega(x) = \sum_a \int_{\mathcal{C}_a} dy \frac{\rho(y)}{x - y} \ . \tag{4.7}$$

The discontinuity of $\omega(x)$ across a point $x \in \mathcal{C}_a$ gives the density:

$$\omega(x + \epsilon) - \omega(x - \epsilon) = 2\pi i \rho(x) \ , \qquad x \in \mathcal{C}_a \ , \tag{4.8}$$

where $\epsilon$ is a suitable infinitesimal.

The saddle-point equation expresses the zero force condition on a test eigenvalue in the presence of the distribution of a large number eigenvalues along each of the cut. It can be written succinctly in terms of the resolvant as

$$\frac{1}{S} V'(x) = 2\mathbf{P}\omega(x) - e^{i\beta}\omega(e^{i\beta}x) - e^{-i\beta}\omega(e^{-i\beta}x) \ ; \qquad x \in \mathcal{C}_a \ , \tag{4.9}$$

for $a = 1, \ldots, N$, where $\mathbf{P}$ implies a principal value, in other words an average of $\omega(x)$ just above and below the cut at $x$:

$$\mathbf{P}\omega(x) = \frac{1}{2}(\omega(x + \epsilon) + \omega(x - \epsilon)) \ , \tag{4.10}$$

where $\epsilon$ is a suitable infinitesimal.

The content of this saddle point equation becomes more transparent when recast in terms of a new function $t(x)$ (see also [16]) defined by

$$t(x) = f(x) + Sx\big(e^{-i\beta}\omega(e^{-i\beta}x) - \omega(x)\big) \ , \tag{4.11}$$

where $f(x)$ is a polynomial defined by

$$f(x) - f(xe^{i\beta}) = xV'(x) \ . \tag{4.12}$$

¿From the analytic structure of the resolvent $\omega(x)$ it follows that $t(x)$ has cuts along each $\mathcal{C}_a$ and its rotation $\mathcal{C}'_a = e^{i\beta}\mathcal{C}_a$. The saddle-point equation (4.9) is then very simple:

$$\mathbf{P}t(x) = \mathbf{P}t(e^{i\beta}x) \ , \qquad x \in \bigcup_a \mathcal{C}_a \ . \tag{4.13}$$

Given (4.8) this is simply a gluing condition which glues the top/bottom of $\mathcal{C}_a$ to the bottom/top of $e^{i\beta}\mathcal{C}_a$

$$t(x \pm \epsilon) = t(e^{i\beta}(x \mp \epsilon)) \ , \qquad x \in \bigcup_a \mathcal{C}_a \ . \tag{4.14}$$

So $t$ defines a Riemann surface $\Sigma_N$ of genus $N$ which is a copy of the complex $x$-plane with the cuts identified as above. The function $t$ is then the unique meromorphic function on $\Sigma_N$ with a pole at $x = \infty$ of the form

$$t(x) \underset{|x|\to\infty}{\longrightarrow} f(x) + \mathcal{O}(1/x) \ . \tag{4.15}$$

On the contrary, $x$ is multi-valued on $\Sigma_N$. If we define a basis of 1-cycles on $\Sigma_N$ $\{A_a, B_a\}$, where $A_a$ encircles $\mathcal{C}_a$ and $B_a$ joins a point $x \in \mathcal{C}_a$ to its image $e^{i\beta}x \in \mathcal{C}'_a$—and hence is a closed cycle—then $x$ is single-valued around each $A_a$ but picks up a multiplicative factor $e^{i\beta}$ around each $B_a$. It appears that the Riemann surface $\Sigma_N$ has $2N$ complex moduli, given by the positions of the ends of the $N$ cuts $\mathcal{C}_a$. However, the fact that there must exist a function $t$ on $\Sigma_N$ with the prescribed singularity of order $N$ at $x = \infty$, as in (4.15), means that there are actually only $N$ moduli.[6]

The $N$ moduli of the surface are encoded in the quantities $S_a = g_s\hat{N}_a$ which can be expressed as the following contour integrals:

$$S_a = S\int_{\mathcal{C}_a} \rho(x)\,dx = -\frac{1}{2\pi}\oint_{A_a} \frac{t\,dx}{x} \ . \tag{4.16}$$

[6]The argument relies on the Riemann-Roch Theorem. Firstly, $t$ has a pole of order $N$ at $v = \infty$. There are $2N$ such functions. However the singular part of $t$ is fixed (4.15) giving $N$ conditions. So the net number of remaining moduli are $2N - N = N$ as claimed.

The free-energy of the matrix model around the saddle-point solution $F(S_a) = \log Z$, which is a funciton of the moduli of of the solution $\{S_a\}$, has the usual topological genus expansion

$$F(S_a) = \sum_{g=0}^{\infty} F_g(S_a) g_s^{2g-2} . \tag{4.17}$$

The quantum vacuum of the field theory is described by an effective superpotential which is a function of the $S_a$ which are now interpreted as the glueball superfields:

$$W_{\text{eff}}(S_a) = \sum_{a=1}^{N} \Big( \frac{\partial F_0}{\partial S_a} - 2\pi i \tau S_a \Big) , \tag{4.18}$$

where $\tau$ is the complexified coupling of the supersymmetric gauge theory in four dimensions and $F_0(S_i)$ is the genus zero component of the free energy.[7]

We already have an expression for the $S_a$ in terms of an integral of a meromorphic form along 1-cycles of $\Sigma_N$. One can also find a similar expression for the other quantities in (4.18):

$$\frac{\partial F_0}{\partial S_a} = -i \oint_{B_a} \frac{t\, dx}{x} . \tag{4.19}$$

A critical point of $W_{\text{eff}}(S_j)$ corresponds to

$$\sum_{a=1}^{N} \frac{\partial^2 F_0}{\partial S_b \partial S_a} = 2\pi i \tau \qquad b = 1, \ldots, N . \tag{4.20}$$

This equation can be written in a more suggestive way by noticing that

$$\omega_a = -\frac{1}{2\pi} \frac{\partial}{\partial S_a} \frac{t\, dx}{x} , \quad a = 1, \ldots, N \tag{4.21}$$

are a basis for the abelian differentials of the first kind on $\Sigma_N$ normalized by

$$\oint_{A_a} \omega_b = \delta_{ab} . \tag{4.22}$$

The reason is that the singular part of $t\, dx/x$ at $x = \infty$ depends only on $V(x)$ and so is manifestly independent of the moduli $S_a$. Taking the $S_b$ derivative of (4.16) then proves the result. Hence,

$$\frac{\partial^2 F_0}{\partial S_b \partial S_a} = 2\pi i \oint_{B_a} \omega_b = 2\pi i \tau_{ab} , \tag{4.23}$$

[7]Note that for the vacuum in question the eigenvalues are non-degenerate, $N_a = 1$. In more general vacua the expression for the glueball superpotential involves the degeneracy [16].

where $\tau_{ab}$ are elements of the period matrix of $\Sigma_N$. Consequently the critical point equations are

$$\sum_{a=1}^{N} \tau_{ab} = \tau \qquad b = 1, \ldots, N \ . \tag{4.24}$$

Given that the moduli space of $\Sigma_N$ is $N$-dimensional, these $N$ conditions completely fix the geometry of the Riemann surface $\Sigma_N$ in terms of the parameters of the probe potential $V(x)$.

## 4.1 Identification of the critical Riemann surface

The curve $\Sigma_N$ at the critical point of the glueball superpotential defines the Seiberg-Witten curve of the $U(N)$ theory at a point on the Coulomb branch determined by the probe potential $V(x)$, $x_a = \xi_a$.

The condition (4.24) implies that $\Sigma_N$ at the critical point of the glueball superpotential is an $N$-fold cover of the torus $E(\tau)$ with complex structure $\tau$. We can cover $E(\tau)$ with a coordinate $z$ defined modulo $2\omega_1 = 2\pi i$ and $2\omega_2 = 2\pi i\tau$ with $\tau = \omega_2/\omega_1$. The covering map $z(P):\ \Sigma_N \to E(\tau)$ is then

$$z(P) = 2\pi i \int_{P_0}^{P} \sum_{a=1}^{N} \omega_a \mod 2\pi i, 2\pi i\tau \ , \tag{4.25}$$

where $P_0$ is a fixed, but otherwise arbitrary, base point.

The fact that the curve $\Sigma_N$ is an $N$-fold cover of $E(\tau)$ means that it can be described by an equation of the form $F(z, x) = 0$ which depends implicitly on the form of $V(x)$; in other words, on the position on the Coulomb branch $\mathcal{C}_1$. Since $\Sigma_N$ covers $E(\tau)$ $N$ times, the function $F(z, x)$ must be of the form

$$F(z, x) = \sum_{a=0}^{N} f_a(z)x^a = \prod_{a=1}^{N} (x - x_a(z)) = 0 \ , \tag{4.26}$$

where we can choose $f_N(z) = 1$. The function $F(z, x)$ must satisfy the further conditions:

**(i)** Quasi-periodicity in $z$. Notice that $x$ is single-valued around the $A_a$ cycles but multi-valued around the $B_a$ cycles (which are lifts of the $A$ and $B$ cycles on $E(\tau)$) we have

$$F(z + 2\pi i, x) = F(z, x) \ , \qquad F(z + 2\pi i\tau, e^{-i\beta}x) = F(z, x) \ . \tag{4.27}$$

In other words, the coefficient functions are quasi-elliptic: $f_a(z+2\pi i) = f(z)$ and $f_a(z+2\pi i\tau) = e^{ia\beta} f_a(z)$. In terms of the $N$ roots $x_a(z)$ this condition becomes

$$x_a(z+2\pi i) = \Sigma^{(1)}_{ab} x_b(z) \ , \qquad x_a(z+2\pi i\tau) = e^{i\beta}\Sigma^{(2)}_{ab} x_b(z) \ , \tag{4.28}$$

where $\Sigma^{(1)}$ and $\Sigma^{(2)}$ are elements of the permutation group $S_N$.

**(ii)** Recall that $\Sigma_N$ is a copy of the $x$-plane with cuts identified in pairs. Hence, $x$ should have a single simple pole on $\Sigma_N$ corresponding to the point at infinity. Hence, exactly one of the $N$ roots $x_a(z)$ should have a simple pole on the torus. Apart from this, the roots $x_a(z)$ should have no other singularities in the period parallelogram. By choosing a suitable origin for $z$ we can arrnge this singularity to sit over $z = 0$ in the cover. Consequently $F(z,x)$ behaves near $z = 0$ as

$$F(z,x) \sim \frac{g(x)}{z} + \mathcal{O}(z^0) \ , \tag{4.29}$$

where $g(x)$ is a polynomial in $x$ of degree at most $N-1$, and has no other singularities in the period parallelogram.

As mentioned above, condition **(i)** means that the coefficient functions $f_a(z)$ are quasi-elliptic. On the other hand, condition **(ii)** constrains these functions to have at most a simple pole in each period parallelogram. As in the previous Section, one may use an argument based on the Riemann-Roch theorem, to count the number of independent complex functions satisfying these conditions. In this way one finds that the total number of moduli of the most general curve satisfying the consitions **(i)** and **(ii)** is $N$ as expected.

# 5 Finding the curve

## 5.1 The 5d theory

Our strategy for finding the family of complex curves $\Sigma_N$ which satisfies conditions **(i)** and **(ii)**, is based on the observation that exactly the same conditions arise in the solution of a completely different $U(N)$ gauge theory. The theory in question arises from the compactification of a supersymmetric gauge theory in $4+1$ dimensions on a circle of radius $R$. The theory has minimal supersymmetry in $4+1$ dimensions (8 supercharges) which reduces to $\mathcal{N}=2$ supersymmetry after compactification to $3+1$ dimensions. This theory (to be called simply "the 5d theory" in the following) contains a $U(N)$ vector multiplet and a massive adjoint hypermultiplet. In $4+1$ dimensions the hypermultiplet mass is a real parameter $m$. An additional 'twisted' mass parameter $\mu$ arises after compactification. The twisted mass

corresponds to the Wilson line of a background gauge field around the compact direction and has the periodicity $\mu \sim \mu + 2\pi$. The theory has a dimensional gauge coupling $G_5^2$. At energies far below the compactification scale the 5d theory reduces to a four-dimensional gauge theory with coupling $G_4^2 = G_5^2/2\pi R$. A four dimensional vacuum angle $\Theta$ can also be introduced by including appropriate couplings to background fields. It is convenient to define the complex combination,

$$M = m + \frac{i\mu}{2\pi R} \tag{5.1}$$

which can be regarded as the lowest component of a background vector multiplet of $\mathcal{N} = 2$ SUSY in four dimensions.

In $4+1$ dimensions, the vector multiplet contains a real adjoint scalar field $\varphi$. The five-dimensional gauge theory described above has a Coulomb branch parametrized by the eigenvalues of this field. An additional adjoint scalar $\omega = \int_{S^1} A \cdot dx$ arises from the Wilson line of the $4+1$-dimensional gauge field around the compact direction. The four-dimensional low-energy theory has $\mathcal{N} = 2$ supersymmetry and includes a vector multiplet whose lowest component is the complex adjoint scalar,

$$\Phi = \varphi + \frac{i\omega}{2\pi R} \tag{5.2}$$

The Coulomb branch is parametrised in terms of the eigenvalues of $\Phi$,

$$\langle \Phi \rangle = \mathrm{diag}(\rho_1, \rho_2, \ldots, \rho_N) \tag{5.3}$$

Apart from the usual action of Weyl gauge transformations which permute the eigenevalues, the periodicity of the Wilson line implies the gauge identification: $\rho_a \sim \rho_a + i/R$ for $a = 1, 2, \ldots, N$. Alternatively we can work in terms of the gauge invariant moduli,

$$U_n = \frac{1}{N}\langle \mathrm{Tr}_N \left[\exp\left(2n\pi R\Phi\right)\right]\rangle = \frac{1}{N}\sum_{a=1}^{N} \exp\left(2n\pi R\rho_a\right) \tag{5.4}$$

for $n = 1, 2, \ldots, N$.

The low energy theory on the Coulomb branch is a four dimensional $U(1)^N$ gauge theory with $\mathcal{N} = 2$ supersymmetry. In the quantum theory, the low-energy action depends on a matrix $\tau_{ab}^{\mathrm{5d}}$ of complexified abelian couplings which varies as a function of the Coulomb branch moduli $U_n$. In the exact solution of the system presented in [17], $\tau_{ab}^{\mathrm{5d}}$ is identified is identified with the period matrix of the spectral curve $\Sigma_N$ of the $N$-body Ruijsenaars-Schneider (RS) integrable system. We will now explain why this curve naturally provides a general solution to conditions equivalent to (**i**) and (**ii**). Our strategy will be to realise the classical 5d theory on an intersection of branes in Type IIA string theory. Following [19], we then obtain the quantum corrections to the Coulomb branch by lifting to M-theory.

## 5.2 The IIA brane configuration

We begin by considering IIB string theory on $\mathbf{R}^{8,1} \times \mathbf{S}^1$ with coordinates $x_0, x_1, \ldots, x_9$. The compact direction is parametrized by the coordinate $x_6$ with $x_6 \sim x_6 + 2\pi R_6$. We introduce $N$ coincident D5 branes with world-volume in the $\{0,1,2,3,4,6\}$ directions. At energies far below the string scale $M_s = 1/\sqrt{\alpha'}$, the worldvolume theory is $\mathcal{N} = (1,1)$ SUSY gauge theory with gauge group $U(N)$ defined on $\mathbf{R}^{4,1} \times \mathbf{S}^1$ with six-dimensional gauge coupling $G_6^2 = 16\pi^3\alpha' g_s$. At energies below the compactification scale $1/R_6 \ll M_s$, this in turn reduces to maximally supersymmetric $U(N)$ gauge theory on $\mathbf{R}^{4,1}$ with gauge coupling $G_5^2 = G_6^2/2\pi R_6$.

We now introduce an additional compact direction via the identification $x_4 \sim x_4 + 2\pi R$ with $R \gg R_6$. The low-energy theory at scales far below $1/R_6$ is then the maximally supersymmetric $4+1$ dimensional theory formulated on $\mathbf{R}^{3,1} \times \mathbf{S}^1$. An equivalent brane configuration which gives rise to the same low energy theory is obtained by performing a T-duality transformation in the $x_4$ direction. This yields a configuration of $N$ D4 branes in Type IIA string theory. The IIA spacetime is $\mathbf{R}^{7,1} \times \mathbf{S}^1 \times \mathbf{S}^1$ where the compact coordinates are $x_4$ and $x_6$ with radii $R_4 = \alpha'/R$ and $R_6$. The D-branes are wrapped on the $x_6$ circle as before but are located at a point in the $x_4$ direction.

Separating the D4 branes in the compact $x_4$ direction corresponds to turning on a Wilson line for the $(4+1)$-dimensional gauge field. More generally, the world-volume theory of the D4 branes has a moduli space corresponding to the motion of the branes in their transverse directions. In terms of minimal supersymmetry in five dimensions (which has eight supercharges), the theory which lives on the branes includes a $U(N)$ vector multiplet and a single massless hypermultiplet. As mentioned above, the vector mutiplet includes a real adjoint scalar $\varphi$ and, after compactification, the Wilson line $\omega$ provides an additional scalar. We will focus on configurations where these fields have non-trivial VEVs as described in (5.3) above. These are realised by seperating the D4 branes in the $x_4$ and $x_5$ directions. In particular we define a complex coordinate $u = (ix_4 + x_5)/R_4$ and place the $D4$ branes at positions $u = u_a$ for $a = 1, 2, \ldots, N$ in the complex $u$-plane. Comparing the spectrum of open strings stretched between the D4 branes with the gauge theory spectrum of W-bosons shows that we can identify the positions in terms of the complex eigenvalues appearing in (5.3). Explicitly, we have $u_a = \rho_a(2\pi\alpha')/R_4$ for $a = 1, 2, \ldots, N$. Note that the periodicity of $u$ (which is $u \sim u + 2\pi i$) matches that of the eigenvalues $\rho_a$ (namely $\rho \sim \rho + i/R$) by virtue of the relation $R_4 = \alpha'/R$.

Finally, to obtain the 5d theory of interest we must introduce a complex mass $M$ for the adjoint hypermultiplet. This can be accomplished by following the same proceedure used in [19] to introduce a hypermultiplet mass in the corresponding four-dimensional the-

ory (the $\mathcal{N} = 2^*$ theory). To do this we introduce an NS5 brane with world-volume in the $\{0, 1, 2, 3, 4, 5\}$ directions. This by itself has no effect on the low energy world-volume theory. To introduce the mass $M$, we include a twist in the boundary conditions in the $x_6$ direction. Specifically, rather than simply compactifying the $x_6$ direction, we divide out by the transformation,

$$x_6 \to x_6 + 2\pi R_6 \ , \qquad u \to u + 2\pi RM \ . \tag{5.5}$$

This twist forces each D4 brane to break, so that its two endpoints on the NS5 brane are no longer coincident but are seperated by a distance $(2\pi\alpha')|M|$ in the complex $u$-plane. The hypermultiplet degrees of freedom correspond to strings stretched between endpoints of D4 branes on either side of the NS5 and hence they will acquire non-zero masses $|M|$ as required.

## 5.3 Lifting to M-theory

To obtain the curve controlling the Coulomb branch of the 5d theory we will lift the IIA brane configuration described above to M-theory. Neglecting the twist in the $x_6$ direction, the IIA spacetime in which the branes live has the form $\mathbf{R}^{7,1} \times \mathbf{S}^1 \times \mathbf{S}^1$. By IIA/M duality, this is equivalent to M-theory on $\mathbf{R}^{7,1} \times \mathbf{S}^1 \times \mathbf{S}^1 \times \mathbf{S}^1$ where the additional compact direction, with coordinate $x_{10}$, has radius $R_{10}$. The M-theory parameters are related to the IIA string lengthscale and coupling as,

$$R_{10} = \sqrt{\alpha'} g_s \ , \qquad M_{\rm Pl} = \frac{g_s^{-1/3}}{\sqrt{\alpha'}} \ . \tag{5.6}$$

Here $M_{\rm Pl}$ denotes the eleven dimensional Planck mass.

There are two further refinements of the standard IIA/M duality we will need. The first is to introduce a non-trivial vacuum angle $\Theta$ in the low-energy theory on the branes and the second is to reintroduce the hypermultiplet masses via the twist (5.5). In fact both these modifications can be incorporated by slanting the torus in the M-theory geometry appropriately. It is convenient to work in terms of dimensionless complex variables $u = (ix_4 + x_5)/R_4$ and $z = (-x_6 + ix_{10})/R_{10}$. The resulting M-theory spacetime can be thought of as $\mathbf{R}^{6,1} \times \mathcal{M}_{\mathbf{C}}$ where $\mathcal{M}_{\mathbf{C}}$ is a two dimensional complex manifold with coordinates $u$ and $z$. The complex manifold in question is obtained as a quotient,

$$\mathcal{M}_{\mathbf{C}} = \frac{\mathbf{C} \times \mathbf{C}}{\Gamma_1 \times \Gamma_2 \times \Gamma_3} \tag{5.7}$$

where $\Gamma_1$, $\Gamma_2$ and $\Gamma_3$ are the complex translations,

$$\begin{aligned} \Gamma_1 :& \quad z \to z + 2\omega_1 \\ \Gamma_2 :& \quad z \to z + 2\omega_2 \qquad u \to u + 2\pi MR \\ \Gamma_3 :& \quad u \to u + 2\pi i \end{aligned} \tag{5.8}$$

with $\omega_1 = i\pi$ and $\omega_2 = i\pi\tau$ where

$$\tau = \frac{iR_6}{R_{10}} + \frac{\Theta}{2\pi} = \frac{4\pi i}{G_4^2} + \frac{\Theta}{2\pi} \tag{5.9}$$

In the case $M = 0$, dividing out by first two translations provides the standard definition of the flat complex torus with complex structure $\tau$, as defined in Section 4.1, and the manifold $\mathcal{M}_{\mathbf{C}}$ is simply $E(\tau) \times \mathbf{C}$. On reintroducing a non-zero hypermultiplet mass $M$, this space is no longer a Cartesian product, but should be thought of as a non-trivial complex line bundle over $E(\tau)$ [19].

Our starting point was a configuration of $N$ D4 branes and a single NS5 brane with intersection $\mathbf{R}^{3,1}$. Both types of IIA brane lift to M-theory fivebranes. As usual we expect to find a single M5 brane with worldvolume $\mathbf{R}^{3,1} \times \hat{\Sigma}_N$ where $\hat{\Sigma}_N$ is a Riemann surface of genus $N$ embedded as a complex submanifold of $\mathcal{M}_{\mathbf{C}}$. As we start from $N$ D4 branes wrapped on the compact $x_6$ direction, the resulting M5 brane will wrap the torus $E(\tau)$, N times. The corresponding Riemann surface will therefore be a branched $N$-fold cover of $E(\tau)$, which can be described as an $N^{\rm th}$ order polynomial in the variables $z$ and $x = \exp(-u)$,

$$F(z, x) = f_N(z) \prod_{a=1}^{N} (x - x_a(z)) \tag{5.10}$$

which is automatically invariant under $\Gamma_3$. For invariance under $\Gamma_1$ and $\Gamma_2$ we require,

$$F(z + 2\omega_1, x) = F(z, x) \ , \qquad F(z + 2\omega_2, e^{2\pi MR}x) = F(z, x) \ . \tag{5.11}$$

This matches precisely with condition (**i**) described in the previous Section provided we identify,

$$\tau = \frac{4\pi i}{G_4^2} + \frac{\Theta}{2\pi} \ , \qquad \beta = -2\pi i M R \ . \tag{5.12}$$

An additional condition on $F(z, x)$ comes from considering the allowed zeros and poles of the roots $x_a(z)$ on the torus $E(\tau)$. The poles correspond to points on the torus where the $M5$ brane goes to infinity in the complex $u$-plane. This is the expected behaviour at points where the original IIA brane configuration extends to infinity in the $x_4$ and $x_5$ directions. This occurs only at the position of the NS5 brane, which corresponds to the point $z = 0$. Hence, exactly one of the $N$ roots $x_a(z)$ should have a simple pole at $z = 0$. Apart from this, the roots $x_a(z)$ should have no other singularities in the period parallelogram. This is equivalent to condition **(ii)** described in the previous section.

In summary the Riemann surface $\hat{\Sigma}_N$ controlling the Coulomb branch of the 5d theory obeys exactly the same conditions as those which constrain the Riemann surface $\Sigma_N$ which

plays the same role for the Coulomb branch $\mathcal{C}_1$ in the $\beta$-deformed theory. Fortunately, the Riemann surface $\hat{\Sigma}_N$ is has been determined independently by Nekrasov [17] to be the spectral curve of the $N$-body Ruijsenaars-Schneider (RS) integrable system. We will check momentarily that the curve indeed provides the unique solution to the conditions **(i)** and **(ii)** given above. Thus our conclusion is that the desired complex curve $\Sigma_N$ must also be the RS spectral curve.

## 5.4 The curve

The spectral curve of the RS system is given explicitly as,

$$F(z,x) = \det\left(L(z) - x\mathbf{1}_{(N)}\right) = 0 \ , \tag{5.13}$$

where $L(z)$ is the $N \times N$ Lax matrix of the RS integrable system with elements

$$L_{ab}(z) = i\varrho_a \frac{\sigma(q_{ab} - i\beta + z)}{\sigma(x_{ab} - i\beta)\sigma(z)} e^{\xi(i\pi)\beta z/\pi} \ , \tag{5.14}$$

where $q_{ab} = q_a - q_b$ and

$$\varrho_a = e^{p_a} \prod_{b(\neq a)} \sqrt{\wp(q_{ab}) - \wp(i\beta)} \ . \tag{5.15}$$

Here $\wp(z)$, $\sigma(z)$ and $\xi(z)$ are standard Weierstrass functions for the torus $E(\tau)$.

The first task is to show that $F(z,x)$ as defined in (5.13) satisfies conditions (**i**) and (**ii**) given above in Section 4.1. This is accomplished in Appendix A. We will now discuss some of the features of the curve $\Sigma_N$.

The complex parameters $q_a$ and $p_a$, for $a = 1, 2, \ldots, N$ correspond to the positions and momenta respectively of $N$ particles. The spectral curve only depends on these variables through the $N$ conserved Hamiltonians;

$$H_n = \sum_{1\leq a_1 \leq \cdots \leq a_n \leq N} \prod_{i=1}^{n} \varrho_{a_i} \prod_{1\leq i<j\leq n} \frac{1}{\wp(i\beta) - \wp(q_{a_i a_j})} \ , \tag{5.16}$$

for $n = 1, 2, \ldots, N$. As explained in the previous section this is the expected number of moduli for the most general solution of conditions **(i)** and **(ii)**. As the Coulomb branch $\mathcal{C}_1$ is parametrized by the $N$ complex moduli $u_n$, we should find some relation between these quantities and the Hamiltonians $H_n$. This relation is constrained by the non-anomalous R-symmetry $U(1)_R^{(1)}$ which acts non-trivially on $\mathcal{C}_1$. The modulus $u_n$ has charge $n$ under this symmetry. On the other hand the curve has an obvious symmetry under which $v \to e^{i\alpha} v$,

$\varrho_a \to e^{i\alpha}\varrho_a$, under which $H_n$ has charge $n$. Identifying these symmetries provides a constraint on the relation between $u_n$ and $H_n$ but does not fix it uniquely. For example, $u_2$ might be identified with any linear combination of $H_2$ and $H_1^2$ where the two coefficients can depend on the couplings $\tau$ and $\beta$. A comparison with one-loop perturbation theory will at least allow us to fix this ambiguity in the weak coupling limit $\tau \to i\infty$.

In fact, the perturbative limit of the period matrix $\tau_{ab}$ of $\Sigma_N$ was calculated in the context of the five-dimensional theory described above [20, 21]. The result reads

$$\tau_{ab} \sim \delta_{ab}\tau + \frac{i}{2\pi}(1-\delta_{ab})\log\left[\frac{\sinh^2 2\pi R(\rho_a-\rho_b)}{\sinh 2\pi R(\rho_a-\rho_b+M)\sinh 2\pi R(\rho_a-\rho_b-M)}\right] , \tag{5.17}$$

where $\rho_a$ are the eigenvalues defined in (5.3) above. The is agrees precisely with the sum of our classical and one-loop results (3.3) provided we identify $x_a = \exp(2\pi R\rho_a)$.

In fact we can do much better than this and show by direct calculation that the relation between the abelian couplings of the two theories described above persists to all orders in the instanton expansion. Details of this calculation, which uses localisation techniques to calculate the instanton contributions directly are given in Appendix B. In summary we find that,

$$\tau_{ab}^{(5d)}(\rho_a, M) = \tau_{ab}\big(x_a = e^{2\pi R\rho_a}, \beta = -2i\pi RM\big) . \tag{5.18}$$

which confirms the equivalence of the two Coulomb branch theories described above.

The curve $\Sigma_N$ has interesting quasi-modular properties under $SL(2,\mathbf{Z})$ transformations acting on the microscopic coupling constant $\tau$. The modular group acts as,

$$\tau \to \tilde{\tau} = \frac{a\tau+b}{c\tau+d} , \tag{5.19}$$

for integers $a$, $b$, $c$ and $d$ with $ad-bc=1$. Under this transformation a function $f(z|\tau)$ defined on the torus $E(\tau)$ has holomorphic modular weight $w$ if,

$$f(z|\tilde{\tau}) = (c\tau+d)^w f(z(c\tau+d)|\tau) . \tag{5.20}$$

With this definition the function the Weierstrass function $\wp(z)$ has weight $+2$. The quasi-elliptic functions $\xi(z)$ and $\sigma(z)$ have weights $+1$ and $-1$ respectively. Thus we find that the equation $F(z,x)=0$ is modular invariant if we assign weights $-1$, $0$ to $q_a$, $\varrho_a$ and weights $-1$, $+1$ to $\beta$ and $x$ respectively.

At $\beta=0$, the modular group acting on $\tau$ corresponds to the exact S-duality of the $\mathcal{N}=4$ theory. The modular invariance of the curve suggests that S-duality extends for non-zero $\beta$, provided we assign a holomorphic modular weight of $-1$ to the deformation parameter.

Thus S-duality transformations relate theories with different values of $\beta$. The reason why we should expect such a duality in the $\beta$-deformed theory was explained in [3, 7]. At linear order, the $\beta$ deformation of the $\mathcal{N}=4$ theory corresponds to adding a (SUSY descendent of a) chiral primary operator $\hat{\mathcal{O}}$ to the $\mathcal{N}=4$ Lagrangian with coupling $\beta$. As the operator $\hat{\mathcal{O}}$ has known modular weight $+1$, modular invariance of the $\mathcal{N}=4$ theory can be restored by assigning the coupling $\beta$ weight $-1$

### 5.5 The multiple branch structure

In the earlier parts of Section 5, we have argued that the curve $\Sigma_N$ describing the Coulomb sub-branches $\mathcal{C}_i$ is the spectral curve of the $N$-body RS integrable system. This matches the result of the instanton analysis. However, the instanton analysis goes much further in that it describes all the Coulomb sub-branches $\mathcal{C}_{n_1,n_2,n_3}$. We can now see how all the sub-branches are described in the language of the integrable system. First of all, the roots of the multiple branches occur when one of the eigenvalues $x_a \to 0$. It is easy to see that this corresponds in the integrable system to the associated momenta $p_a \to -\infty$ or $\varrho_a \to 0$. In this limit, the $N$-body RS system naturally degenerates to the $N-1$-body RS system. Notice that these points of degeneration do not occur in the moduli space of the five-dimensional theory since they would require $\rho_a = -\infty$. As one moves out onto the branch $\mathcal{C}_{n_1,n_2,n_3}$, the associated integrable system consists of 3 non-interacting copies of the RS system with $n_1$, $n_2$ and $n_3$ particles. So this Coulomb sub-branch is holomorphically equivalent to a five-dimensional gauge theory with product gauge group $U(n_1)\times U(n_2)\times U(n_3)$. This equivalence is confirmed by the explicit instanton calculations described in Appendix B.

## 6 Explicit Results for Gauge Group $U(2)$

For $N=2$, the spectral curve of the RS integrable system is,

$$F(z,x) = x^2 - H_1 f_\beta(z) x + H_2 f_\beta^2(z)\big(\wp(i\beta) - \wp(z-i\beta)\big) = 0 \ , \tag{6.1}$$

where

$$f_\beta(z) = \frac{\sigma(z-i\beta)}{\sigma(-i\beta)\sigma(z)} e^{\xi(i\pi)\beta z/\pi} \ . \tag{6.2}$$

The two Hamiltonians are given as,

$$H_1 = i(e^{p_1} + e^{p_2})\sqrt{\wp(q_1-q_2) - \wp(i\beta)} \ , \qquad H_2 = e^{p_1+p_2} \ . \tag{6.3}$$

Defining a new variables $t = xf_\beta(z)/\sqrt{H_2} - \mathcal{U}$, with $\mathcal{U} = H_1/2\sqrt{H_2}$, the curve takes on the simpler form,

$$t^2 = \mathcal{U}^2 - \wp(i\beta) + \wp(z - i\beta) \ . \tag{6.4}$$

In this form the curve $\Sigma_2$ is a manifestly a double cover of the standard complex torus $E(\tau)$ (with periods $2\omega_1 = 2\pi i$ and $2\omega_2 = 2\pi i\tau$). Invariance under the modular group acting on $\tau$ is manifest if we assign $\beta$ holomorphic modular weight $-1$ as above and $t$ and $\mathcal{U}$ both have modular weight $+1$. An interesting double periodicity in $\beta$ is also apparant. In particular, note that the theory is obviously invariant under shifts of $\beta$ by integer multiples $2\omega_1/i = 2\pi$. This is because the classical superpotential (1.1) is invariant under this shift up to an overall change sign which can be absorbed by redefining the fields. However, the curve is also invariant under shifts of $\beta$ by multiples of $2\omega_2/i = 2\pi\tau$. This periodicity is not visible in the classical theory and the period itself is non-perturbative in the gauge coupling.

For generic values of $z \in E(\tau)$, the quadratic (6.4) has two distinct roots

$$t_\pm = \pm\sqrt{\mathcal{U}^2 - \wp(i\beta) + \wp(z - i\beta)} \ , \tag{6.5}$$

the branch points of the double cover, occur at special values of $z$ for which the two roots coincide: $t_+ = t_- = 0$. This occurs for values of $z$ satisfying,

$$\wp(z - i\beta) = \wp(i\beta) - \mathcal{U}^2 \tag{6.6}$$

As $\wp(z)$ is an elliptic function of order two, it attains each complex value exactly twice in each period parallelogram. More precisely $\wp(z - i\beta) - u$, considered as a function of $z$, has exactly two simple zeros for each value of $u$, excepting the three special values $u = e_i(\tau)$, $i = 1, 2, 3$, for which the function has one double zero. Thus, for generic values of $\mathcal{U}$, there are two distinct values, $z_1$ and $z_2$, which satisfy (6.6). These are the two branch points of the double-covering, and the two sheets of the covering are joined along a cut which runs from $z_1$ to $z_2$.

As usual we are interested in finding special points in the moduli space parametrized by $\mathcal{U}$, where the curve degenerates. This happens when the two branch points $z_1$ and $z_2$ coincide up to periods of $E(\tau)$. From the above discussion, this happens only when the RHS of equation (6.6) attains one of the three special values $e_i(\tau)$. Thus we find a total of six critical points in the moduli space for which

$$\mathcal{U} = \mathcal{U}^{(i)}_\pm = \pm\sqrt{\wp(i\beta) - e_i(\tau)} \tag{6.7}$$

for $i = 1, 2, 3$. At these points the curve necessarily degenerates to an unbranched double cover of the bare torus $E(\tau)$.

To understand the significance of these points we will compute the period matrix $\tau_{ab}$ of the curve which yields the low-energy abelian gauge couplings of the $\beta$-deformed $U(2)$ theory.

We must first find a convenient basis for the homology of $\Sigma$. To do this we define a canonical set of cycles on the surface $\{A_a, B_a\}$, with $a, b = \pm$, and intersections $A_a \cdot B_a = \delta_{ab}$. $A_\pm$ ($B_\pm$) are simply the lift of the $A$ and $B$ cycles on the torus $E(\tau)$ (corrsesponding to $z \sim z + 2\pi i$ and $z \sim z + 2\pi i\tau$, respectively) to the two sheets $t = t_\pm$, where the cycles are chosen to avoid the cut in the $z$ plane. As we are dealing with a surface of genus two, the space of holomorphic differentials is two dimensional. A convenient basis is,

$$d\Omega_1 = dz \ , \qquad d\Omega_2 = dz/t \ . \tag{6.8}$$

We then define the matrices,

$$h_{a\alpha} = \oint_{B_a} d\Omega_\alpha \ , \qquad e_{a\alpha} = \oint_{A_a} d\Omega_\alpha \tag{6.9}$$

in terms of which the period matrix is computed as the matrix product

$$\tau_{ab} = h_{a\alpha} \left(e^{-1}\right)^{\alpha}{}_{b} \tag{6.10}$$

Evaluating this we find,

$$\tau_{11} = \tau_{22} = \tfrac{1}{2}(\tau + \tau_{\rm odd}) \ , \qquad \tau_{12} = \tau_{21} = \tfrac{1}{2}(\tau - \tau_{\rm odd}) \ , \tag{6.11}$$

where $\tau_{\rm odd}$ (defined in Section 3.1) is

$$\tau_{\rm odd} = \frac{\oint_{B_+} d\Omega_2}{\oint_{A_+} d\Omega_2} = \frac{\int_{z_0}^{z_0+\omega_2} \frac{dz}{\sqrt{\mathcal{U}^2 - \wp(i\beta) + \wp(z - i\beta)}}}{\int_{z_0}^{z_0+\omega_1} \frac{dz}{\sqrt{\mathcal{U}^2 - \wp(i\beta) + \wp(z - i\beta)}}} = \frac{-iK'(k)}{K(k)} \ , \tag{6.12}$$

where $z_0$ is arbitrary and $K$ and $K'$ are complete elliptic integrals of the first kind with parameter

$$k = \sqrt{\frac{(\wp(i\beta) - \mathcal{U}^2 - e_1)(e_2 - e_3)}{(\wp(i\beta) - \mathcal{U}^2 - e_2)(e_1 - e_3)}} \ . \tag{6.13}$$

The two periods $\tau_{\rm even} \equiv \tau$ and $\tau_{\rm odd}$ are even and odd respectively under the $\mathbf{Z}_2$ symmetry $t \to -t$ which interchanges the two sheets of $\Sigma_2$. In field theory language, this $\mathbf{Z}_2$ is precisely the Weyl subgroup of $U(2)$. We can compare the extact formula for $\tau_{\rm odd}$ to the one-loop result (3.6) by taking the semiclassical limit $\tau \to i\infty$. In this limit the Weierstrass function reduces to a trigonometric function as,

$$\wp(z) \sim \frac{1}{4\sinh^2\left(\frac{z}{2}\right)} + \frac{1}{12} \ , \tag{6.14}$$

while the quasi-modular forms $e_1$, $e_2$ and $e_3$, tend to the constant values $-1/6$, $1/12$ and $1/12$ respectively. Using standard results for the asymptotics of the complete elliptic integrals we obtain,

$$\tau_{\rm odd} \sim \tau + \frac{1}{i\pi} \log\left(\frac{4\mathcal{U}^2 + \sin^{-2}(\beta/2)}{1 - 4\mathcal{U}^2 - \sin^{-2}(\beta/2)}\right) \quad . \tag{6.15}$$

This matches the perturbative result (3.6), provided that we set

$$\mathcal{U} = \frac{H_1}{2\sqrt{H_2}} = \frac{i\varphi}{2\sin(\beta/2)} = \frac{i}{4\sin(\beta/2)}\left(\sqrt{\frac{x_1}{x_2}} + \sqrt{\frac{x_2}{x_1}}\right) \quad , \tag{6.16}$$

where $\varphi$ was defined in (3.7).

We can also look at the behaviour of $\tau_{\rm odd}$ near the six points in moduli space where the curve degenerates. Near the points $\mathcal{U}^{(1)}_{\pm}$, defined in (6.7) above, we find,

$$\tau_{\rm odd} \sim -\frac{1}{i\pi} \log\left(\mathcal{U} - \mathcal{U}^{(1)}_{\pm}\right) + \ldots \tag{6.17}$$

The coefficient in front of the logarithm is consistent with the appearance of a single massless hypermultiplet electrically charged under $U(1)_{\rm odd}$. We can confirm this interpretation by noting that, in the semiclassical limit $\tau \to i\infty$, we have, $\mathcal{U}^{(1)}_{\pm} \simeq \pm i\cot(\beta/2)/2$. Using the semiclassical identification (6.16) this corresponds the relations $\lambda_1 = \exp(\pm i\beta)\lambda_2$ between the two eigenvalues of $\Phi_1$. These are precisely the points where massless charged hypermultiplets appear in the classical theory.

We next consider the neighbourhood critical points,[8] $\mathcal{U}^{(3)}_{\pm}$. Near either of these points the abelian couplings exhibit the asymptotics,

$$\tau_{\rm odd} \sim -\frac{i\pi}{\log\left(\mathcal{U} - \mathcal{U}^{(3)}_{\pm}\right)} + \ldots \tag{6.18}$$

Thus the low-energy gauge-coupling $g^2_{\rm odd} = 4\pi/{\rm Im}\tau_{\rm odd}$ has a logarithmic divergence near these points. To interpret this we recall the electric-magnetic duality of the low-energy action which, for $N = 2$ includes an $SL(2,\mathbf{Z})$ acting on $\tau_{\rm odd}$. After performing the duality transformation $\tau_{\rm odd} \to -1/\tau_{\rm odd}$, the asymptotics of the dual coupling precisely match those expected for a massless hypermultiplet (*i.e.* (6.17)). The intepretation is therefore that we have a massless hypermultiplet at each of the points $\mathcal{U}^{(3)}_{\pm}$ which is minimally coupled to the dual gauge-field, in other words these degrees of freedom carry magnetic charge.

[8] For pedagogical reasons it is convenient to discuss the critical points in the order $\mathcal{U}^{(1)}_{\pm}$, then $\mathcal{U}^{(3)}_{\pm}$ then $\mathcal{U}^{(2)}_{\pm}$.

Finally in the vicinity of the third critical point $\mathcal{U}_{\pm}^{(2)}$ the behaviour of the coupling is,

$$\tau_{\rm odd} \sim -\frac{i\pi}{\log\left(\mathcal{U}-\mathcal{U}_{\pm}^{(2)}\right)} \; -1+\ldots \tag{6.19}$$

Similar reasoning shows that dyonic hypermultiplets carrying both electric and magnetic charges become massless at these points.

In summary, we have found six critical points in the moduli space. At each of these points an additional massless hypermultiplet appears. The points come in three pairs with massless electric, magnetic and dyonic degrees of freedom respectively. As we vary the deformation parameter $\beta$, we can find special values at which one pair of these points coincide. Here we will focus on the cases where the massless degrees of freedom are mutually local. Thus we look for values of $\beta$ for which $\mathcal{U}_{+}^{(i)}=\mathcal{U}_{-}^{(i)}=0$ for exactly one value of $i$. This requires $\wp(i\beta)=e_i(\tau)$. This is satisfied when $i\beta$ equals one of the the half periods $\omega_1$, $\omega_1+\omega_2$ or $\omega_2$. The three special values are therefore:

**(1)** For $\beta=\pi$ we have $\wp(i\beta)=e_1(\tau)$ which implies $\mathcal{U}_{+}^{(1)}=\mathcal{U}_{-}^{(1)}=0$. Thus, for this special value of $\beta$, two electrically charged hypermultiplets appear at the point $\mathcal{U}=0$. Near the point $\mathcal{U}=0$ we therefore have four light electrically charged chiral multiplets $Q_i$ and $\tilde{Q}_i$ for $i=1,2$. To have these multiplets become massless at the point $\mathcal{U}=0$, we must have a superpotential of the form

$$W \sim x\left(Q_1\tilde{Q}_1+Q_2\tilde{Q}_2\right)\;, \tag{6.20}$$

where $x$ is some local coordinate on the Coulomb branch which goes to zero at the singular point. As explained in Section 2, such an effective theory has a Higgs branch on which the charged chiral multiplets condense breaking the gauge group down to its $U(1)$ center. In fact this is just the Higgs branch (2.13) which is already visible in the classical theory. Our analysis confirms that the Higgs branch is still present in the quantum theory and intersects the Coulomb branch at $\mathcal{U}=0$ as in the classical theory.

**(2)** For $\beta=\pi\tau$ we have $\wp(i\beta)=e_3(\tau)$ which implies $\mathcal{U}_{+}^{(3)}=\mathcal{U}_{-}^{(3)}=0$. For this value of $\beta$, the monodromies are consistent with the existence of four massless magnetically charged chiral multiplets $Q_i^{(M)}$ and $\tilde{Q}_i^{(M)}$, with $i=1,2$, at the point $\mathcal{U}=0$. In addition to magnetic gauge couplings, the effective theory of the light degrees of freedom has a superpotential term of the form

$$W \sim x^{(M)}\left(Q_1^{(M)}\tilde{Q}_1^{(M)}+Q_2^{(M)}\tilde{Q}_2^{(M)}\right)\;, \tag{6.21}$$

where $x^{(M)}$ is a local coordinate on the Coulomb branch which vanishes at the singular point $\mathcal{U}=0$. This effective theory has a three complex dimensional branch on which the magnetically charged fields have non-zero VEVs. Thus we find a branch on which magnetic

states condense leading to the confinement of electric charges. This is one of the confining branches discussed in [3].

**(3)** For $\beta = \pi(\tau + 1)$ we have $\wp(i\beta) = e_2(\tau)$ which implies $\mathcal{U}_+^{(2)} = \mathcal{U}_-^{(2)} = 0$. states carrying both electric and magnetic charges at the point $\mathcal{U} = 0$. The effective theory near the singular point has a branch on which these dyonic degrees of freedom condense. This theory branch on this branch therefore features oblique confinement.

It is interesting to look at the form of the curve which appears at the critical values of $\beta$ identified above. We will work in terms of the original form of the $U(2)$ curve (6.1). In each case the critical value of the modulus $\mathcal{U} = 0$ corresponds to $H_1 = 0$, with arbitrary $H_2$. On this locus the curve becomes;

$$x^2 = H_2 f_\beta^2(z) \left(\wp(z - i\beta) - \wp(i\beta)\right) \tag{6.22}$$

At the critical value $\beta = \pi$ corresponding to the root of the Higgs branch the curve simplifies dramatically and becomes, $x^2 = H_2 F_1(\tau)$ where $F_1(\tau) = 2e_1^2(\tau) + e_2(\tau)e_3(\tau)$ is a quasi-modular form of weight four. The two roots $x_\pm(z)$ are equal to the constants $\pm\sqrt{H_2 F_1(\tau)}$. Under translation by the two periods of the torus the roots behave as,

$$x_\pm(z + 2\omega_1) = x_\pm(z) \ , \qquad x_\pm(z + 2\omega_2) = e^{i\beta} x_\mp(z) \ . \tag{6.23}$$

Thus the minimal half-periods $\tilde{\omega}_i$ such that $x_\pm(z + 2\tilde{\omega}_i) = x_\pm(z)$ for $i = 1, 2$ are $\tilde{\omega}_1 = \omega_1$ and $\tilde{\omega}_2 = 2\omega_2$. The corresponding critical curve is an unbranched double cover of the torus $E(\tau)$ with modular parameter $\tilde{\tau} = 2\tau$.

It is also instructive to look more closely at the behaviour of the curve as we approach the critical point. Thus we set $\beta = \pi + \epsilon$ and expand (6.22) to first order in $\epsilon$,

$$x^2 = H_2 F_1(\tau) + \epsilon \mathcal{A}(z) \tag{6.24}$$

Although the second term is subleading in $\epsilon$, it becomes large near the point $z = 0$. In fact we find $\mathcal{A}(z) \sim c/z$ plus finite terms as $z \to 0$, where $c = -i\wp''(i\pi)$. Near the origin (6.24) becomes,

$$z(x - x_+)(x - x_-) = c\epsilon \tag{6.25}$$

Thus we see that, as $\epsilon \to 0$, the curve actually factorizes into two branches. One branch, defined by $x = x_\pm$ with arbitrary $z$ corresponds to the double cover of the torus $E(\tau)$ discussed above. The new branch is defined by $z = 0$ with arbitrary $x$. Factorization of the curve into two pieces is the usual signal of the appearance of a new branch: in this case the Higgs branch discussed above. In the next section we will make this observation precise in the context of the M-theory construction of the curve.

The behaviour of the curve at the other two critical points is related to the behaviour at the Higgs branch root via S-duality. For $\beta = \pi\tau$ and $H_1 = 0$, the curve becomes $x^2 = H_3F_3(\tau)\exp(\pi iz/\omega_1)$ with $F_3(\tau) = 2e_3^2+e_1e_2$. The two roots are $v_\pm(z) = \pm\sqrt{H_2F_3(\tau)}\exp(\pi iz/2\omega_1)$. Under translation by the two periods of the torus the roots behave as,

$$x_\pm(z+2\omega_1) = x_\mp(z) \ , \qquad x_\pm(z+2\omega_2) = e^{i\beta}x_\pm(z) \ . \tag{6.26}$$

Thus the minimal half-periods $\tilde{\omega}_i$ such that $x_\pm(z+2\tilde{\omega}_i) = x_\pm(z)$ for $i = 1, 2$ are $\tilde{\omega}_1 = 2\omega_1$ and $\tilde{\omega}_2 = \omega_2$. The critical curve is an therefore an unbranched double cover of the torus $E(\tau)$ with modular parameter $\tilde{\tau} = \tau/2$.

Finally, for $\beta = \pi(\tau+1)$ and $H_1 = 0$, the curve becomes $x^2 = H_2F_2(\tau)\exp(\pi iz/\omega_1)$ with $F_2(\tau) = 2e_2^2+e_1e_3$. The two roots are $x_\pm(z) = \pm\sqrt{H_2F_2(\tau)}\exp(\pi iz/2\omega_1)$. Under translation by the two periods of the torus the roots behave as,

$$x_\pm(z+2\omega_1) = x_\mp(z) \ , \qquad x_\pm(z+2\omega_2) = e^{i\beta}x_\mp(z)$$

Thus the minimal half-periods $\tilde{\omega}_i$ such that $x_\pm(z+2\tilde{\omega}_i) = x_\pm(z)$ for $i = 1, 2$ are $\tilde{\omega}_1 = 2\omega_1$ and $\tilde{\omega}_2 = \omega_1+\omega_2$. Thus the critical curve is an unbranched double cover of the torus $E(\tau)$ with modular parameter $\tilde{\tau} = (\tau+1)/2$.

In each case the resulting surface is interpreted as an unbranched double cover of the torus $E(\tau)$ which is itself a torus $E(\tilde{\tau})$ with complex structure parameter $\tilde{\tau} = \tilde{\omega}_2/\tilde{\omega}_1$. Thus in the three degenerate cases $\beta = \pi$, $\pi\tau$ and $\pi(\tau+1)$ we found $\tilde{\tau} = 2\tau$, $\tau/2$ and $(\tau+1)/2$ respectively. These values correspond to the three inequivalent unbranched double covers of $E(\tau)$ and they are naturally permuted by S-duality.

The generalisation of these results to gauge group $U(N)$ with $N > 2$ is straightforward. As explained in Section 2, the classical theory has a Higgs branch on which the unbroken gauge group is $U(1)$ for $\beta = 2\pi/N$. This branch survives in the quantum theory and corresponds to a degeneration of $\Sigma_N$ to an unbranched $N$-fold cover of the torus $E(\tau)$ with complex structure $\tilde{\tau} = N\tau$. In the quantum theory, there are also branches in confining phases which are related to the Higgs branch by S-duality. These correspond to the degenerations of $\Sigma_N$ into inequivalent $N$-fold covers of $E(\tau)$. These correspond to all torii with complex structure $\tilde{\tau} = (p\tau+k)/q$ where $pq = N$ and $k = 0, 1, \ldots, q-1$. Thus the total number of inequivalent branches is equal to the sum of the divisors on $N$. Note that this is essentially identical to the classification of massive vacua of the $\mathcal{N} = 1^*$ theory [22–24]. In the present case, each branch occurs at a set of critical values of the form $\beta = 2\pi(l+m\tilde{\tau})/p$ where $l$ and $m$ are integers.

# 7 Discussion

An interesting consequence of the analysis given above is that two very different supersymmetric gauge theories have Coulomb branches described by the same complex curve $\Sigma_N$. The $\beta$-deformed theory, which is the main subject of the paper, lives in $3+1$ dimensions, has four supercharges and also has spontaneously broken conformal invariance. The other theory (the 5d theory of Section 5 above) lives in $4+1$ dimensions, has eight supercharges and no conformal invariance. Despite these differences, our results imply that, on their Coulomb branches these two models agree exactly at the level of $\mathcal{N}=1$ F-terms. Indeed, below Eq (5.17), we identified the holomorphic change of variables which relates the two Coulomb branch theories at one loop: it is simply the exponential map $x_a = \exp(2\pi R\rho_a)$ for $a = 1, 2, \ldots, N$. In Appendix B, we check that this relation continues to hold order by order in the instanton expansion and therefore for the exact matrix of abelian couplings. In this Section, we will check that the pattern of Higgs and Confining branches in the two theories also agrees. Our conclusion therefore is that the theories are actually holomorphically equivalent in the sense explained in Section 1. In this Section we will also discuss some of the consequences of this equivalence for the 5d theory.

As the curves for the two theories agree, the Coulomb branches of both models have the same singular points, with equivalent monodromies and massless states. The 5d theory should therefore exhibit the same Higgs and Confining branches intersecting the Coulomb branch at the singularities. The Higgs branch roots of the 5d theory are straightforward to find. As the 5d theory has eight supercharges the stable states of the theory are BPS saturated. The classical mass spectrum of elementary quanta on the 5d Coulomb branch is governed by the central charge,

$$\mathcal{Z}_{ab} = \frac{ik}{R} + \rho_a - \rho_b \pm M \tag{7.1}$$

Here the integer $k$ corresponds to momentum around the compact $x_4$ direction which has radius $R$. The complex eigenvalues $\rho_a$, for $a = 1, 2, \ldots, N$ are defined in Section 5.1 above as is the complex mass $M$. A Higgs branch of the $U(N)$ $\beta$-deformed theory occurs at $\beta = 2\pi/N$. According to (5.12), the equivalent value of the 5d parameter $MR$ is $i/N$. If we also specify the Coulomb branch VEVs as $\rho_a = ai/NR$ for $a = 1, 2, \ldots, N$ the central charge becomes,

$$\mathcal{Z}_{ab} = \frac{i}{NR}\,(kN + a - b \pm 1) \tag{7.2}$$

Thus we find a total of $N$ massless $\mathcal{N}=2$ hypermultiplets coming from off-diagonal elements of $\mathcal{Z}$ with[9] $a = b \pm 1 \bmod N$. Each of these fields carry opposite charges under a pair of

[9]Each $\mathcal{N}=2$ hypermultiplet includes two $\mathcal{N}=1$ chiral multiplets of opposite charges.

adjacent $U(1)$'s in the Cartan subalgebra of $U(N)$. This precisely the same massless spectrum which appears at the root of the corresponding Higgs branch of the $\beta$ deformed theory.

The presence of the Higgs branch in the classical theory can be seen directly in the IIA brane construction of Section 5.2. The values for the complex eigenvalues $\rho_a$ described above correspond to a configuration where the $N$ D4 branes are distributed around the compact $x_4$ direction with equal spacings $2\pi R_4/N$. At the special value $MR = i/N$ the twist (5.5) in the $x_6$ direction becomes,

$$x_6 \to x_6 + 2\pi R_6 \ , \qquad x_4 \to x_4 + \frac{2\pi R_4}{N} \ . \tag{7.3}$$

As the resulting shift in $x_4$ is equal to the seperation between branes, the $N$ D4 branes segments form a closed spiral. In other words they form a single D4 wrapped $N$ times around the $x_6$ direction and once around the $x_4$ direction. This D4 can now move away from the NS5 in the $x_7$, $x_8$ and $x_9$ directions. As we now have a single D4 brane, this corresponds to the expected Higgs branch where $U(N)$ is broken down to $U(1)$. The D4 can also move parallel to the NS5 in the $x_4$ and $x_5$ directions and we we may turn on a Wilson line for the $U(1)$ world-volume gauge field around the $x_6$ circle. Thus the Higgs branch has a total of six real or three complex dimensions as expected. One may easily check that other Higgs branches which occur when multiple spirals can be formed match the remaining classical Higgs branches of the $\beta$-deformed theory.

The brane picture of the Higgs branch root can easily be lifted to M-theory. The D4 brane described above lifts to a single M5 with two wrapped worldvolume dimensions. The fivebrane is wrapped $N$ times on the torus $E(\tau)$ in the M-theory spacetime with no branch-points or other singularities. Thus the M5 world-volume has the form $\mathbf{R}^{3,1} \times \Sigma$ where $\Sigma$ is an unbranched N-fold cover of $E(\tau)$. As the M5 winds $N$ times round the $x_6$ direction and once round the $x_{10}$ direction the relevant $N$-fold covering has complex structure $\tilde{\tau} = N\tau$. The NS5 brane lifts to a second M5 brane located at $z = 0$ and infinitely extended in the $x_4$ and $x_5$ directions. In the case $N = 2$, this configuration of two intersecting M5 branes precisely corresponds to the factorization curve described in the previous section. As above, the Higgs branch corresponds to moving the two branes apart in their common transverse dimensions.

Another interesting aspect of the results presented above was their invariance under $SL(2, \mathbf{Z})$ transformations acting on the complexified coupling constant $\tau$ and on the deformation parameter $\beta = -2\pi i RM$. In Section 5.4, we explained the significance of this duality in the context of the $\beta$-deformed theory. It is also of interest to understand this duality in the context of the 5d theory. In the undeformed case, $M = 0$, we are considering the maximally supersymmetric $U(N)$ gauge theory on $\mathbf{R}^{3,1} \times \mathbf{S}^1$. At energies far below the compactification scale $1/R$, this reduces to $\mathcal{N} = 4$ SUSY Yang-Mills in four dimensions with

complexified coupling $\tau$, and the $SL(2,\mathbf{Z})$ in question is simply the usual S-duality of this theory. However this is not only a duality of the low-energy theory. In fact, as we now review, it is an exact duality of the full theory on $\mathbf{R}^{3,1}\times\mathbf{S}^1$.

Recall that the maximally supersymmetric theory in five dimensions is itself equivalent to a compactification of the $(2,0)$ superconformal theory which lives in six dimensions. Specifically the 5d theory with $M=0$ (and $\Theta=0$) corresponds to a compactification of the $A_{N-1}$ $(2,0)$ theory[10] on $\mathbf{R}^{3,1}\times\mathbf{S}^1\times\tilde{\mathbf{S}}^1$. The first compact dimension with radius $R$ is already apparent in the 5d theory. The second circle has radius set by the five-dimensional gauge coupling: $\tilde{R}=G_5^2/8\pi^2$. From the six-dimensional viewpoint, the electric-magnetic duality transformation $\tau\to-1/\tau$ (with $\Theta=0$) simply corresponds to an interchange of the two compact dimensions: $R\leftrightarrow\tilde{R}$. This implies the exact equivalence of the 5d theory with coupling $G_5^2$ and radius of compactification $R$, with a dual theory with coupling $\tilde{G}_5^2=8\pi^2R$ and compactification radius $\tilde{R}$. More generally, we may introduce a four-dimensional vacuum angle $\Theta$ by replacing $\mathbf{S}^1\times\tilde{\mathbf{S}}^1$ with a slanted torus of complex structure $\tau$. The $SL(2,\mathbf{Z})$ duality group acting on $\tau$ corresponds to the diffeomorphism group of this torus.

Our results, and those of [17], indicate that this duality extends to the 5d theory with a non-zero hyper-multiplet mass. In particular the dimensionless combination $MR$, like the deformation parameter $\beta$, transforms with modular weight $(-1,0)$ under this duality. A related point is that $MR$ like $\beta$ has two periods.

$$MR\to MR+i\ ,\qquad MR\to MR+i\tau\ . \tag{7.4}$$

The first period is already apparent in the classical theory and is a consequence of the identification of the twisted mass $\mu=2\pi R\mathrm{Im}[M]$ as a background Wilson line. The second period (as noted in [17]) is non-perturbative in the coupling and is therefore invisible in the classical theory. It would be interesting to understand these features from the point of view of the six-dimensional $(2,0)$ theory.

One of the most interesting features of the $\beta$ deformed theory is the existence of branches of confining vacua corresponding to the condensation of magnetic monopoles. The existence of these branches was first argued in [3] as a consequence of S-duality acting on the Higgs branches of the theory and, in the previous section, we checked this explicitly for gauge group $U(2)$. Having established the existence of both Higgs branches and S-duality in the 5d theory it naturally follows that corresponding confining branches are also present in this theory. It is not hard to identify the massless states which condense on these branches. Recall that, in addition to elementary quanta, the 5d theory contains various solitonic states. In addition to ordinary four-dimensional magnetic monopoles which are independent of the compact

[10]To produce a $U(N)$ gauge theory in five dimensions we should also include an additional free tensor multiplet in the six dimensional theory.

spatial coordinate, the theory contains solitons corresponding to Yang-Mills instantons on $\mathbf{R}^3\times\mathbf{S}^1$ treated as static solutions of finite energy on $\mathbf{R}^{3,1}\times\mathbf{S}^1$. There are also boundstates of these objects which carry both instanton number and magnetic charge. Each of these states saturates a BPS bound and their classical masses are determined by the central charge,

$$\tilde{\mathcal{Z}}_{ab} = \frac{iq}{\tilde{R}} + (\rho_a - \rho_b)\frac{i\tilde{R}}{R} \pm M \tag{7.5}$$

where $\tilde{R} = G_5^2/R$ as above. Here $q$ corresponds to the instanton number while the contribution proportional to $\rho_a - \rho_b$ comes from states magnetically charged under the corresponding Cartan $U(1)$ subgroup. The central charge (7.5) is evidently S-dual to (7.1)[11].

The confining branch root occurs (for $\Theta = 0$) at the point $MR = 8\pi^2 i/G_5^2 N = iR/N\tilde{R}$. As at the Higgs branch root, the eigenvalues of the adjoint scalar take the values $\rho_a = ai/NR$ for $a = 1, 2, \ldots, N$. Hence the central charge (7.5) becomes

$$\tilde{\mathcal{Z}}_{ab} = \frac{i}{N\tilde{R}}\,(qN + a - b \pm 1) \tag{7.6}$$

Thus we find a total of $N$ massless $\mathcal{N} = 2$ hypermultiplets coming from off-diagonal elements of $\tilde{\mathcal{Z}}$ with $a = b \pm 1$ mod $N$. Each of these fields carry opposite magnetic charges under a pair of adjacent $U(1)$'s in the Cartan subalgebra of $U(N)$. This precisely the same massless spectrum which appears at the root of the corresponding confining branch of the $\beta$ deformed theory.

The presence of the new branch can also be seen directly in the M-theory construction of Section 5.3. Near the root of the branch, the curve factorizes into an M5 brane wrapped $N$ times on the torus $E(\tau)$ corresponding to the compact $x_6$ and $x_{10}$ directions and a single flat M5 brane extending in the $x_4$ and $x_5$ directions. This configuration is related to the configuration at the Higgs branch root by a $6 - 10$ flip. Correspondingly the first M5 brane winds once around the $x_6$ direction and $N$ times around the $x_{10}$ direction to give an $N$-fold cover of $E(\tau)$ with complex structure $\tilde{\tau} = \tau/N$. The existence of these branches of vacua was also noted in the string theory construction of [29].

The new branches of the 5d theory described above exhibit interesting physics. We find the confinement of a $U(N)$ gauge group down to its decoupled central $U(1)$ due to monopole condensation coexisting with unbroken $\mathcal{N} = 2$ supersymmetry (in the four dimensional

[11]We emphasize that both (7.5) and (7.1) are classical formulae for the central charge. In a theory with $\mathcal{N} = 2$ supersymmetry, we generally expect that the central charges and the BPS mass spectrum will recieve quantum corrections. Interestingly, in the present case, the classical formulae yield the exact results for the location of the Higgs and confining branches (ie they agree with our earlier calculation based on the curve $\Sigma_N$). Thus, for these special values of the parameters and moduli, it seems that the classical formulae (7.1) and (7.5) are actually exact.

sense). As in the $\beta$-deformed theory confinement occurs without a mass gap, although the interpretation of the resulting massless scalars is different. In the $\beta$-deformed theory, the massless scalar fields could be interpreted as Goldstone bosons for the spontaneously broken scale invariance and R-symmetry. In the 5d case the massless scalars correspond to the traces of adjoint scalar fields and are completely decoupled.

The holomorphic equivalence described above only applies to the Coulomb branch $\mathcal{C}_1 = \mathcal{C}_{N,0,0}$ of the $\beta$-deformed theory and the Higgs and confining branches which intersect it. It is straightforward to extend these results to the generic Coulomb branch $\mathcal{C}_{n_1,n_2,n_3}$ with $n_1 + n_2 + n_3 = N$. This branch is described by a complex curve of the form $\Sigma_{n_1} \bigcup \Sigma_{n_2} \bigcup \Sigma_{n_3}$ which is in turn appropriate to govern the vacuum structure of the 5d theory with gauge group $U(n_1) \times U(n_2) \times U(n_3)$.

We can also extend our discussion to the $\beta$-deformed theory with gauge group $SU(N)$. On the Coulomb sub-branch $\mathcal{C}_1$ and at the level of the integrable system, the traceless constaint appropriate for the $SU(N)$ theory corresponds to imposing the tracelessness of the Lax operator. This is equivalent to imposing $H_1 = 0$.[12]. The proof of this simple condition is somewhat involved and we have relegated it to Appendix C. Notice that the constraint is different from the constraint that gives the five-dimensional $SU(N)$ theory from the five-dimensional $U(N)$ theory. In that case, the $U(1)$ factor corresponds to the centre-of-mass motion of the integrable system, so the constraint is $\sum_a p_a = 0$, or $H_N = 1$. This means that the holomorphic equivalence of the two theories is not valid when the gauge group is $SU(N)$. On one of the more general Coulomb sub-branches one simply imposes $H_1$ in each of the three copies of the RS integrable system.

In closing we note that the holomorphic equivalence described in this section suggests a direct way of realising the $\beta$-deformed theory on the world volume of Type IIA branes. The set up involves $N$ D4 branes wrapped around a compact dimension $x_6 \sim x_6 + 2\pi R_6$ and extended in the $\{0, 1, 2, 3\}$ directions. We will also include a single NS5 brane extended in the $\{0, 1, 2, 3, 4, 5\}$ directions. We define a complex coordinate $U = x_4 + ix_5$. So far this is simply a four dimensional version of the construction of subsection 5.2. The low energy theory on the branes (at energy scales much less than $1/R_6$) is just $\mathcal{N} = 4$ SUSY Yang-Mills with gauge group $U(N)$. As in eqn (5.5) of subsection 5.2, an adjoint hypermultiplet mass can be introduced by introducing an additive shift in $U$ on going around the $x_6$ circle. This leads to the standard brane construction of the $\mathcal{N} = 2^*$ theory in four dimensions given in [19]. Instead of doing this, we will introduce a *multiplicative* twist in $U$ via the identification.

$$x_6 \to x_6 + 2\pi R_6 \ , \qquad U \to \exp(i\beta) U \ . \tag{7.7}$$

It is straightforward to verify that the resulting spectrum of stretched strings reproduces

[12]Note that imposing $H_1 = Nm$ corrsponds to the $SU(N)$ theory with the mass term $m\mathrm{Tr}\,\Phi_2\Phi_3$ added to the superpotential

the classical spectrum of the $\beta$-deformed theory on its Coulomb branch $\mathcal{C}_1$. One may also show that lifting this configuration to M-theory correctly reproduces the curve $\Sigma_N$ given in Section 5 above. It is tempting to conclude that the twist (7.7) has the effect introducing the $\beta$-deformation of the $\mathcal{N}=4$ theory on the worldvolume of the branes. However, this cannot be quite correct as the other Coulomb branches of the $\beta$-deformed theory are not visible in this brane construction. Equivalently, the construction only produces the correct deformation of vacua in the appropriate region of the Coulomb branch of the $\mathcal{N}=4$ theory. It would be interesting to understand this in more detail.

ND acknowledges useful discussions with Ofer Aharony. ND would like to thank the organizers of the 'QCD and String Theory' workshop at the ITP Santa Barbara where part of this work was completed. TJH would like to thank all the organizers and participants of the 2nd Simons Workshop on Mathematical Physics in Stony Brook for providing a stimulating atmosphere for completing this work.

## Appendix A: Properties of the Curve

To check property (**i**) we need the quasi-periodic properties of the function $\sigma(z)$:

$$\sigma(z+2\omega_1) = -\sigma(z)\exp\left(2(z+\omega_1)\xi(\omega_1)\right) \ , \quad \sigma(z+2\omega_2) = -\sigma(z)\exp\left(2(z+\omega_2)\xi(\omega_2)\right) \quad \text{(A.1)}$$

and the relation $\omega_2\xi(\omega_1) - \omega_1\xi(\omega_2) = i\pi/2$. Recall in our conventions $2\omega_1 = 2\pi i$ and $2\omega_2 = 2\pi i\tau$. Using these relations we find that,

$$L_{ab}(z+2\pi i) = \left(U_1 L(z) U_1^{-1}\right)_{ab} \ , \qquad L_{ab}(z+2\pi i\tau) = e^{i\beta}\left(U_2 L(z) U_2^{-1}\right)_{ab} \ , \qquad \text{(A.2)}$$

where the gauge transformations $U_1$ and $U_2$ are given by,

$$\left(U_1\right)_{ab} = \exp\left(2x_a\xi(\pi i)\right) \ , \qquad \left(U_2\right)_{ab} = \exp\left(2x_a\xi(\pi i\tau)\right) \qquad \text{(A.3)}$$

Using these relations in (5.13) we can easily check that $F(z+2\pi i, x) = F(z,x)$ and $F(z+2\pi i\tau, e^{-i\beta}x) = F(z,x)$ as required by condition **(i)**.

To check condition (**ii**), we need to find the singularities of $F(z,x)$, which arise at singular points of the Lax matrix elements $L_{ab}(z)$ given in (5.14) above. The quasi-elliptic function $\sigma(z)$ behaves as $\sigma(z)\sim z + O(z^5)$ near $z=0$ and has no other zeros or singularities in the period parallelogram. Thus the only singularity of $F(z,x)$ lies at $z=0$. Near $z=0$ we have,

$$L_{ab}(z) \sim \frac{1}{z}L_{ab}^{(-1)} + L_{ab}^{(0)} + \mathcal{O}(z) \ , \qquad \text{(A.4)}$$

where $L^{(-1)}_{ab} = i\rho_a$. As $L^{(-1)}$ is a projection operator onto the vector $(1,1,\ldots,1)$ we can easily change basis so that,

$$\tilde{L}^{(-1)}_{ab} = \big(UL^{(-1)}U^{-1}\big)_{ab} = i\tilde{\rho}_a\delta_{b1} \tag{A.5}$$

for some element $U \in GL(N,\mathbf{C})$. We also define,

$$\tilde{L}^{(0)}_{ab} = \big(UL^{(0)}U^{-1}\big)_{ab} \ . \tag{A.6}$$

As the determinant defining $F(z,v)$ is invariant under this change of basis we can write,

$$\begin{aligned} F(z,x) &= \det\big(L(z) - xI_{(N)}\big) \sim \det\left(\frac{1}{z}\tilde{L}^{(-1)} + \tilde{L}^{(0)} - xI_{(N)}\right) + \mathcal{O}(z) \\ &= \begin{vmatrix} \frac{i}{z}\tilde{Q}_1 - x - \tilde{L}^{(0)}_{11} & \tilde{L}^{(0)}_{12} & \ldots\ldots & \tilde{L}^{(0)}_{1N} \\ i\tilde{Q}_2\frac{1}{z} + \tilde{L}^{(0)}_{21} & -x + \tilde{L}^{(0)}_{22} & \ldots\ldots & \tilde{L}^{(0)}_{2N} \\ \ldots & \ldots & \ldots\ldots & \ldots \\ \ldots & \ldots & \ldots\ldots & \ldots \\ i\tilde{Q}_N + \frac{1}{z}\tilde{L}^{(0)}_{N1} & \ldots & \ldots\ldots & \tilde{L}^{(0)}_{NN} \end{vmatrix} + \mathcal{O}(z) \ . \end{aligned} \tag{A.7}$$

Evaluating the leading term in this determinant we find,

$$F(z,x) \sim \frac{i\tilde{\rho}_1 f(x)}{z} + O(z^0) \tag{A.8}$$

where $f(x)$ is a polynomial in $x$ of order $N-1$. This verifies condition (**ii**).

## Appendix B: The Instanton Calculus

In this appendix, we briefly describe how one can write down the collective coordinate integrals over the moduli space of instantons which determine the instanton contribution to the low-energy couplings $\tau_{ab}$. We also sketch how one can show that the instanton contribution to the couplings $\tau_{ab} = 0$ when $a \in \mathfrak{I}_i$, $b \in \mathfrak{I}_j$ with $i \neq j$.

The instanton caluclus in the $\beta$-deformed theories can be deduced from that of the $\mathcal{N}=4$ theory by taking a careful account of the $\beta$ deformation. We will use the notation and results from the review [25]. At leading order in the semi-classical expansion—which is the level required to calculate the instanton contributions to $F$-term—the effect of the $\beta$ deformation is to modify a subset of the Yukawa couplings. In order to describe the deformation, we first relate the $\mathcal{N}=4$ notation of [25] to the $\mathcal{N}=1$ notation which is appropriate to the situation at hand. In the $\mathcal{N}=4$ theory the relevant Yukawa couplings are of the form

$$\mathrm{Tr}(\lambda^{\alpha A}\bar{\Sigma}_{\hat{a}AB}[\varphi_{\hat{a}},\lambda^B_\alpha]) \ . \tag{B.1}$$

Here, $\varphi_{\hat{a}}$ is an $SO(6)$ vector.[13] First of all, let us relate this to the language of $\mathcal{N}=1$. The three $\mathcal{N}=1$ chiral fields $\Phi_i=(\phi_i,\psi_{i\alpha})$, $i=1,2,3$, are given by

$$\Phi_1=(-\varphi_5+i\varphi_6,\lambda^1_\alpha)\ ,\quad \Phi_2=(\varphi_3-i\varphi_4,\lambda^2_\alpha)\ ,\quad \Phi_3=(-\varphi_1+i\varphi_2,\lambda^3_\alpha)\ , \tag{B.2}$$

so note that $\psi_{i\alpha}\equiv\lambda^i_\alpha$, while $\lambda^4_\alpha\equiv\lambda_\alpha$ is the gluino. The Yukawa couplings (B.1) can then be written in $\mathcal{N}=1$ language as

$$\sum_{ijk}\epsilon_{ijk}\psi^\alpha_i[\phi_j,\psi_{k\alpha}]+\sum_i\psi^\alpha_i[\phi^\dagger_i,\lambda_\alpha]+\text{h.c.}\ . \tag{B.3}$$

The $\beta$-deformation then replaces the commutator in the first term by the deformed commutator. In the language of $\mathcal{N}=4$, this can be achieved by modifying the Clebsch-Gordon coeeficients in the following way:

$$\varphi_a\bar{\Sigma}_{\hat{a}AB}\to\varphi_a\bar{\Sigma}^{(\beta)}_{\hat{a}AB}$$
$$=\begin{pmatrix}0 & e^{i\beta/2}(-\varphi_1+i\varphi_2) & e^{-i\beta/2}(-\varphi_3+i\varphi_4) & -\varphi_5-i\varphi_6\\ e^{-i\beta/2}(\varphi_1-i\varphi_2) & 0 & e^{i\beta/2}(-\varphi_5+i\varphi_6) & \varphi_3+i\varphi_4\\ e^{i\beta/2}(\varphi_3-i\varphi_4) & e^{-i\beta/2}(\varphi_5-i\varphi_6) & 0 & -\varphi_1-i\varphi_2\\ \varphi_5+\varphi_6 & -\varphi_3-i\varphi_4 & \varphi_1+i\varphi_2 & 0\end{pmatrix}\ . \tag{B.4}$$

The fact that the $\beta$ deformation modifies the Yukawa couplings does not affect the structure of the instanton (quasi-)zero modes: at leading-order the gauge field and fermions take the same form in the instanton backgound. The first effect is to change the solution for the scalar fields at leading-order in the instanton background but in a rather simple way. In the $\mathcal{N}=4$ theory, the solution is given in Eq. (4.64) of [25]. One now simply replaces

$$\bar{\Sigma}_{\hat{a}AB}\ \text{by}\ \bar{\Sigma}^{(\beta)}_{\hat{a}AB}\ . \tag{B.5}$$

The same replacement in Eq. (5.25) of [25] gives $\tilde{S}^{(\beta)}$ the *instanton effective action* in the $\beta$ deformed theory. In the $\mathcal{N}=4$ case, by introducing some auxiliary variables one can relate the instanton collective coordinate system to the theory of D-instantons inside D3-brane in Type IIB string theory. We can capture the collective coordinate system of instantons in the $\beta$-deformed theory by simply making the global replacement (B.5). The expression for the instanton action is then

$$\tilde{S}^{(\beta)}=4\pi^2\text{tr}\Big\{\big|w_{\dot{\alpha}}\chi_{\hat{a}}+\varphi^0_{\hat{a}}w_{\dot{\alpha}}\big|^2-[\chi_{\hat{a}},a'_n]^2-\tfrac{1}{2}\bar{\Sigma}^{(\beta)}_{\hat{a}BA}\bar{\mu}^A\varphi^0_{\hat{a}}\mu^B\chi_{\hat{a}}+\bar{\Sigma}^{(\beta)}_{\hat{a}AB}(\bar{\mu}^A\mu^B+\mathcal{M}'^A\mathcal{M}'^B)\chi_{\hat{a}}\Big\}+\tilde{S}_{\text{L.m.}}\ , \tag{B.6}$$

which replaces Eq. (6.94) of [25]. The Lagrange multipler term imposes the bosonic and fermionic ADHM constraints:

$$\tilde{S}_{\text{L.m.}}=-4\pi^2\text{tr}\Big\{\bar{\psi}^{\dot{\alpha}}_A\big(\bar{\mu}^Aw_{\dot{\alpha}}+\bar{w}_{\dot{\alpha}}\mu^A+[\mathcal{M}'^{\alpha A},a'_{\alpha\dot{\alpha}}]\big)+D^c\big(\tau^{c\dot{\alpha}}_{\dot{\beta}}(\bar{w}^{\dot{\beta}}w_{\dot{\alpha}}+\bar{a}'^{\dot{\beta}\alpha}a'_{\alpha\dot{\alpha}})-\zeta^c\big)\Big\}\ . \tag{B.7}$$

[13] In order to avoid confusion with the gauge index, we indicate $SO(6)$ indices as $\hat{a}$.

In the above, the VEVs are the elements of the diagonal $N \times N$ matrix $\varphi^0_{\hat{a}}$ which are given in terms of the $x_a$ by the correspondence (B.2). The quantities $w_{\dot\alpha}$ and $\mu^A$ are $N \times k$ matrices while $\bar{w}^{\dot\alpha}$ and $\bar\mu^A$ are $k \times N$ matrices (with $\bar{w}^{\dot\alpha} \equiv (w_{\dot\alpha})^\dagger$). The remaininf ones, $\mathcal{M}'^A_\alpha$, $a'_{\alpha\dot\alpha}$, $D^c$, $\hat\chi_{\hat a}$ and $\bar\psi^{\dot\alpha}_A$ are all $k \times k$ matrices.

Now we turn to the instanton contributions to the couplings $\tau_{ab}$. These are given by the integrals over the instanton moduli space in (3.15). In $\mathcal{N} = 2$ theories the analogous integrals enjoy certain localization properties [30,31].[14] We now argue that this localization extends to the $\mathcal{N} = 1$ theory. As in the $\mathcal{N} = 2$ theories the argument rests on the existence of a nilpotent fermionic symmetry $Q$. It the $\mathcal{N} = 1$ theory it is simply one of the two supersymmetries that are unbroken by the instanton (taken with a $c$-number parameter). For example we can choose $Q \equiv Q_1$. After the instanton action has been linearized (see Section 6.5 of [25]), it has the structure

$$\tilde{S}^{(\beta)} = Q\Xi + \Gamma \ , \tag{B.8}$$

where $Q\Gamma = 0$ so that $Q\tilde{S}^{(\beta)} = 0$ (up to $U(k)$ transformations). Here,[15]

$$\Xi = -2i\pi^2 \bar\Sigma_{\hat a 4 i} \mathrm{tr}\big(\bar\mu^i (w_1 \chi_{\hat a} + \varphi^0_{\hat a} w_{\dot\alpha}) - (\chi_{\hat a} \bar w^2 - \bar w^2 \varphi^0_{\hat a})\mu^i\big) \tag{B.9}$$

and

$$\begin{aligned} Q\Xi = 4\pi^2 \mathrm{tr}\Big\{ &\big| w_{\dot\alpha}\chi_{\hat a} + \varphi^0_{\hat a} w_{\dot\alpha}\big|^2 + \tfrac12 \bar\Sigma_{\hat a 4 i}(-\bar\mu^4 \varphi^0_{\hat a}\mu^i + \bar\mu^i \varphi^0_{\hat a}\mu^4 + \bar\mu^4\mu^i\chi_{\hat a} - \bar\mu^i\mu^4\chi_{\hat a}) \\ &- \bar\psi^{\dot\alpha}_i(\bar\mu^i w_{\dot\alpha} + w_{\dot\alpha}\mu^i)\Big\} \ . \end{aligned} \tag{B.10}$$

It then follows that if we introduce a coupling $s$, via $\tilde{S}^{(\beta)} \to s^{-1}Q\Xi + \Gamma$, then the resulting integrals which give the coupling cannot depend on $s$. This follows from the fact that the integrals are invariant under the supersymmetry $Q_1$. Taking $s \to \infty$, one can evaluate the integrals around the zeros of $Q\Xi$. These are given by taking the $k \times k$ matrices $\chi_{\hat a}$ to be diagonal with elements which equal one of the non-zero diagonal elements of the VEV matrix $\varphi_{\hat a}$ (there are $N$ such elements $x_a$, $a = 1, \ldots, N$). So the number of critical points is equal to the number of ways of distributing $k$ objects into $N$ sets. Suppose $k \to k_1 + \ldots + k_N$, where we allow $k_a = 0$ for some values of $a$. Each of the $k \times k$ matrix variables in the instanton calculus has a block form which matches this partition: we denote the blocks with the notation $[\cdots]_{ab}$. For example $[\mathcal{M}'^A]_{ab}$. At the critical points one can verify that all the matrix variables are block diagonal. These collective coordinates describe an instanton configuration which consist of $k_a$ abelian instantons in the $a^{\text{th}}$ $U(1)$ subgroup of $U(N)$.[16]

[14] We remark that this localization is more restricted than that used by Nekrasov [32], however, it has the advantage of easily extending to the $\mathcal{N} = 1$ thoery under discussion.

[15] Note that $\bar\Sigma^{(\beta)}_{\hat a 4 i} \equiv \bar\Sigma_{\hat a 4 i} = -\bar\Sigma_{\hat a i 4}$.

[16] Instantons are non-trivial in an abelian theory once $|\zeta^c| > 0$. This deformation is achieved by making the spacetime theory non-commutative. This deformation is not expected to affect the structure of the Coulomb branch [25,33].

One now expands around the critical point and integrates out fluctuations. The leading order expression is then exact because higher-order terms would depend non-trivially on powers of $s$. The leading-order expression is itself independent of $s$ because of cancellations between the bosonic and fermionic fluctuation determinants. We are interested in the VEV dependence of the couplings $\tau_{ab}$, such dependence arises when one integrates out the off-diagonal fluctuations between each pair of blocks $a$ and $b$ ($a \neq b$). There are two generic situations. Firstly, when $a \in \mathfrak{I}_i$ and $b \in \mathfrak{I}_j$ with $i \neq j$. In this case, taking for example $i = 1$ and $j = 2$, the relevant terms for the fermionic fluctuations in the instanton action are of the schematic form

$$[\bar{\psi}]_{ba} \begin{pmatrix} 0 & 0 & -e^{-i\beta/2}x_b & -x_a^* \\ 0 & 0 & -e^{-\beta/2}x_a & x_b^* \\ e^{i\beta/2}x_b & e^{i\beta/2}x_a & 0 & 0 \\ x_a^* & -x_b^* & 0 & 0 \end{pmatrix} [\psi]_{ab} \ , \tag{B.11}$$

where $\psi$ and $\bar{\psi}$ are generic Grassmann collective coordinates.[17] On integrating out these coordinates one gets determinants of the form

$$\big(|x_a|^2 + |x_b|^2\big)^2 \tag{B.12}$$

which does not depend at all on $\beta$. These determinants will then cancel against bosonic determinants in the denominator.[18] The situations with $a, b \in \mathfrak{I}_i$ is very different. Taking $i = 1$, for instance, the fermionic fluctuations in the instanton action are now of the schematic form

$$[\bar{\psi}]_{ba} \begin{pmatrix} 0 & 0 & 0 & x_b^* - x_a^* \\ 0 & 0 & e^{i\beta/2}x_b - e^{-i\beta/2}x_a & 0 \\ 0 & e^{i\beta/2}x_a - e^{-i\beta/2}x_b & 0 & 0 \\ x_a^* - x_b^* & 0 & 0 & 0 \end{pmatrix} [\psi]_{ab} \ . \tag{B.13}$$

Now the determinant gives

$$(e^{i\beta/2}x_a - e^{-i\beta/2}x_b)(e^{i\beta/2}x_b - e^{-i\beta/2}x_a)(x_a^* - x_b^*)^2 \ . \tag{B.14}$$

The compensating bosonic determinant is $\beta$-independent which fixes the overall dependence on the pair $x_a$ and $x_b$ to be

$$\frac{(e^{i\beta/2}x_a - e^{-i\beta/2}x_b)(e^{-i\beta/2}x_a - e^{i\beta/2}x_b)}{(x_a - x_b)^2} \ . \tag{B.15}$$

This is R-symmetry invariant and holomorphic as required. Note that these considerations match the result of perturbation theory as described in Section 3.1.

[17] Either the pairs $(\bar{\mu}^A, \mu^A)$ or $(\mathcal{M}_1'^A, \mathcal{M}_2'^A)$.

[18] This must happen bacause the VEV dependence should disappear in the $\mathcal{N} = 4$ theory at $\beta = 0$.

In fact we can make a stronger statement about the couplings $\tau_{ab}$ with $a \in \mathfrak{I}_i$ and $b \in \mathfrak{I}_j$ with $i \neq j$; namely the couplings vanish. The reason depends on the behaviour of the insertions $\Xi_{a\alpha}$ in instanton integral. At leading around the critical point $\Xi_{a\alpha}$ only depends on the collective coordinates in the $a^{\rm th}$ block. Hence, when $a \in \mathfrak{I}_i$ and $b \in \mathfrak{I}_j$ with $i \neq j$, the two insertions involve the collective coordinates of two different blocks (in another words of abelian instantons in two different $U(1)$ subgroups of the gauge group). Inevitably this means that to get a non-zero answer in the instanton integral entails going beyond leading order around the critical point. Consequently the result must be a non-trivial function of $s$ a dependence which is not allowed. Hence, we conclude that

$$\tau_{ab} = 0 \text{ when } a \in \mathfrak{I}_i \text{ and } b \in \mathfrak{I}_j \text{ with } i \neq j \ . \tag{B.16}$$

On the coulomb branch $\mathcal{C}_i$ we have sketched above how the instanton contributions can only depend on the VEVs through the functions (B.15). This allows us to make a direct connection with the Coulomb branch of the five-dimensional $\mathcal{N} = 2^*$ theory compactified on a circle. In this theory the couplings (or pre-potential) are determined by instantons in much the same way as above except that the collective coordinates can now depend on $x^5$, the periodic coordinate. The instanton action is precisely as for the $\mathcal{N} = 4$ theory, but now there is an integral over $t \equiv x^5$ and there is an additional mass term [25]. Once again the localizations arguments can be made. For us the interesting point concerns the integrals over the fermionic fluctuations. The relevant term in the instanton action mirrors (B.11), but with an integral over $t$ and the addition of mass terms:

$$\int_0^R dt\, [\bar\psi]_{ba}(t) \begin{pmatrix} 0 & 0 & 0 & \rho_b^* - \rho_a^* \\ 0 & 0 & \rho_b - \rho_a + M & 0 \\ 0 & \rho_a - \rho_b + M & 0 & 0 \\ \rho_a^* - \rho_b^* & 0 & 0 & 0 \end{pmatrix} [\psi]_{ab}(t) \ . \tag{B.17}$$

Here, $\rho_a$ are the VEVs of the adjoint scalar. Now on integrating out the fluctuations one has to take account of all the Kaluza-Klein modes around the circle. The end result is

$$\sinh \pi R(\rho_a - \rho_b + M) \sinh \pi R(\rho_b - \rho_a + M) \sinh^2 \pi R(\rho_a^* - \rho_b^*) \ . \tag{B.18}$$

Once again there must be a compensating determinant from the bosonic fluctuations in order to cancel this when $M = 0$ where the theory has $\mathcal{N} = 4$ supersymmetry. So the dependence on the VEVs is through

$$\frac{\sinh 2\pi R(\rho_a - \rho_b + M) \sinh 2\pi R(\rho_a - \rho_b - M)}{\sinh^2 2\pi R(\rho_a - \rho_b)} \ . \tag{B.19}$$

All the remaining parts of the calculation are identical the four-dimensional case. Putting this together with the perturbative contribution as explained in the text, it follows that

the couplings of the four-dimensional $\beta$-deformed theory on the Coulomb branch $\mathcal{C}_i$ and five-dimensional $\mathcal{N}=2^*$ theory are formally related:

$$\tau_{ab}^{(5d)}(\rho_a,M)=\tau_{ab}\big(x_a=e^{2\pi R\rho_a},\beta=-2i\pi RM\big)\ . \tag{B.20}$$

## Appendix C: $U(N)$ and $SU(N)$

One can deduce the coulomb structure of the $SU(N)$ in the following way. Firstly, at the level of the instanton calculus, one simply imposes the tracelessness of the VEVs; in other words

$$\sum_{a\in\mathfrak{I}_i}x_a=0\ , \tag{C.1}$$

for $i=1,2,3$. This is because the instantons lie purely in the non-abelian part of the gauge group. For example on the Coulomb sub-branch $\mathcal{C}_1$ we have $\sum_{a=1}^N x_a=0$. Notice that in the compactified five-dimensional theory the constraint is different; namely $\sum_{a=1}^N\rho_a=0$, i.e. $\prod_{a=1}^N x_a=1$. So the holomorphic equivalence of the $\beta$-deformed theory and the five-dimensional theory is only true for $U(N)$ gauge groups.

One can show that the constraint $\sum_{a=1}^N x_a=0$ becomes the condition $H_1=0$ in the integrable system, or simply the tracelessness of the Lax matrix. In order to see this we need to have the mapping between the moduli $\{x_a\}$ and $\{H_a\}$. This can be extracted from writing the curve of the five-dimensional theory in terms of the moduli $\{\rho_a\}$:

$$\sum_{n=1}^\infty\frac{1}{n!}\Big(\frac{RM}{i}\Big)^n\partial_z^n\theta_1(z/(2i)|\tau)\partial_u^n\prod_{a=1}^N\sinh(u-2\pi R\rho_a)\ . \tag{C.2}$$

Putting $e^{-u}=x$ and $RM=i\beta/(2\pi)$ gives the curve of the $\beta$-deformed theory. The first Hamiltonian $H_1$, or $\mathrm{Tr}\,L(z)$, is proportional to ratio of the coefficients of the $x^{N-1}$ and $x^N$ terms. This gives

$$H_1\propto\sum_{a=1}^N e^{2\pi R\rho_a}=\sum_{a=1}^N x_a\ , \tag{C.3}$$

as expected.

# References


[1] R. G. Leigh and M. J. Strassler, "Exactly marginal operators and duality in four-dimensional $\mathcal{N}=1$ supersymmetric gauge theory," Nucl. Phys. B **447** (1995) 95 [arXiv:hep-th/9503121].

[2] A. Parkes and P. C. West, "Finiteness In Rigid Supersymmetric Theories," Phys. Lett. B **138** (1984) 99.
P. C. West, "The Yukawa Beta Function In N=1 Rigid Supersymmetric Theories," Phys. Lett. B **137** (1984) 371.
D. R. Jones and L. Mezincescu, "The Chiral Anomaly And A Class Of Two Loop Finite Supersymmetric Gauge Theories," Phys. Lett. B **138** (1984) 293.
S. Hamidi, J. Patera and J. H. Schwarz, "Chiral Two Loop Finite Supersymmetric Theories," Phys. Lett. B **141** (1984) 349.
S. Hamidi and J. H. Schwarz, "A Realistic Finite Unified Theory?," Phys. Lett. B **147** (1984) 301.
W. Lucha and H. Neufeld, "Finiteness Of Quantum Field Theories And Supersymmetry," Phys. Lett. B **174** (1986) 186.
W. Lucha and H. Neufeld, "Finite Quantum Field Theories," Phys. Rev. D **34** (1986) 1089.
D. R. Jones, "Coupling Constant Reparametrization And Finite Field Theories," Nucl. Phys. B **277** (1986) 153.
A. V. Ermushev, D. I. Kazakov and O. V. Tarasov, "Finite N=1 Supersymmetric Grand Unified Theories," Nucl. Phys. B **281** (1987) 72.
X. d. Jiang and X. j. Zhou, "A Criterion For Existence Of Finite To All Orders N=1 Sym Theories," Phys. Rev. D **42** (1990) 2109.

[3] N. Dorey, "S-duality, deconstruction and confinement for a marginal deformation of N =4 SUSY Yang-Mills," JHEP **0408** (2004) 043 [arXiv:hep-th/0310117].

[4] N. Dorey, "A new deconstruction of little string theory," JHEP **0407** (2004) 016 [arXiv:hep-th/0406104].

[5] N. Seiberg and E. Witten, "Electric - magnetic duality, monopole condensation, and confinement in N=2 supersymmetric Yang-Mills theory," Nucl. Phys. B **426** (1994) 19 [Erratum-ibid. B **430** (1994) 485] [arXiv:hep-th/9407087].

[6] A. Kapustin, "The Coulomb branch of N = 1 supersymmetric gauge theory with adjoint and fundamental matter," Phys. Lett. B **398**, 104 (1997) [arXiv:hep-th/9611049].
A. Giveon, O. Pelc and E. Rabinovici, "The Coulomb phase in N = 1 gauge theories with an LG-type superpotential," Nucl. Phys. B **499** (1997) 100 [arXiv:hep-th/9701045].
C. Csaki, J. Erlich, D. Z. Freedman and W. Skiba, "N = 1 supersymmetric product group theories in the Coulomb phase," Phys. Rev. D **56** (1997) 5209 [arXiv:hep-th/9704067].

[7] N. Dorey, T. J. Hollowood and S. P. Kumar, "S-duality of the Leigh-Strassler deformation via matrix models," JHEP **0212** (2002) 003 [arXiv:hep-th/0210239].

[8] T. J. Hollowood, “New results from glueball superpotentials and matrix models: The Leigh-Strassler deformation,” JHEP **0304** (2003) 025 [arXiv:hep-th/0212065].

[9] T. Mansson, “Another Leigh-Strassler deformation through the matrix model,” JHEP **0303** (2003) 055 [arXiv:hep-th/0302077].

[10] R. Dijkgraaf and C. Vafa, “A perturbative window into non-perturbative physics,” arXiv:hep-th/0208048.

[11] R. Dijkgraaf and C. Vafa, “On geometry and matrix models”, [arXiv:hep-th/0207106]

[12] R. Dijkgraaf and C. Vafa, “Matrix models, topological strings, and supersymmetric gauge theories”, [arXiv:hep-th/0206255]

[13] F. Cachazo and C. Vafa, “N = 1 and N = 2 geometry from fluxes,” arXiv:hep-th/0206017.

[14] N. Dorey, T. J. Hollowood, S. Prem Kumar and A. Sinkovics, [arXiv:hep-th/0209089].

[15] N. Dorey, T. J. Hollowood, S. P. Kumar and A. Sinkovics, [arXiv:hep-th/0209099].

[16] T. J. Hollowood and S. Prem Kumar, arXiv:hep-th/0407029.

[17] N. Nekrasov, “Five dimensional gauge theories and relativistic integrable systems,” Nucl. Phys. B **531** (1998) 323 [arXiv:hep-th/9609219].

[18] F. Benini, “The Coulomb branch of the Leigh-Strassler deformation and matrix models”, [arXiv:hep-th/0411057].

[19] E. Witten, “Solutions of four-dimensional field theories via M-theory,” Nucl. Phys. B **500**, 3 (1997) [arXiv:hep-th/9703166].

[20] H. W. Braden, A. Marshakov, A. Mironov and A. Morozov, ‘Seiberg-Witten theory for a non-trivial compactification from five to four dimensions,” Phys. Lett. B **448** (1999) 195 [arXiv:hep-th/9812078].

[21] H. W. Braden, A. Marshakov, A. Mironov and A. Morozov, “The Ruijsenaars-Schneider model in the context of Seiberg-Witten theory,” Nucl. Phys. B **558** (1999) 371 [arXiv:hep-th/9902205].

[22] C. Vafa and E. Witten, “A Strong coupling test of S duality,” Nucl. Phys. B **431** (1994) 3 [arXiv:hep-th/9408074].

[23] R. Donagi and E. Witten, “Supersymmetric Yang-Mills Theory And Integrable Systems,” Nucl. Phys. B **460** (1996) 299 [arXiv:hep-th/9510101].

[24] N. Dorey, "An elliptic superpotential for softly broken N = 4 supersymmetric Yang-Mills theory," JHEP **9907** (1999) 021 [arXiv:hep-th/9906011].

[25] N. Dorey, T. J. Hollowood, V. V. Khoze and M. P. Mattis, "The calculus of many instantons," Phys. Rept. **371** (2002) 231 [arXiv:hep-th/0206063].

[26] S. N. Ruijsenaars, Commun. Math. Phys. **115** (1988) 127.

[27] S. N. Ruijsenaars and H. Schneider, Annals Phys. **170** (1986) 370.

[28] S. N. Ruijsenaars, Commun. Math. Phys. **110** (1987) 191.

[29] I. Bena and R. Roiban, "N = 1* in 5 dimensions: Dijkgraaf-Vafa meets Polchinski-Strassler," JHEP **0311** (2003) 001 [arXiv:hep-th/0308013].

[30] T. J. Hollowood, "Testing Seiberg-Witten theory to all orders in the instanton expansion," Nucl. Phys. B **639**, 66 (2002) [arXiv:hep-th/0202197].

[31] T. J. Hollowood, "Calculating the prepotential by localization on the moduli space of instantons," JHEP **0203**, 038 (2002) [arXiv:hep-th/0201075].

[32] N. A. Nekrasov, "Seiberg-Witten prepotential from instanton counting," arXiv:hep-th/0306211.

[33] T. J. Hollowood, V. V. Khoze and G. Travaglini, "Exact results in noncommutative N = 2 supersymmetric gauge theories," JHEP **0105** (2001) 051 [arXiv:hep-th/0102045].